50

COSAS QUE HAY QUE SABER SOBRE

FILOSOFÍA



BEN DUPRÉ

Con frecuencia se ha considerado a la filosofía como la disciplina académica por antonomasia, con sus practicantes firmemente enclaustrados en sus torres de marfil, al margen de los problemas de la vida real. Nada más lejos de la realidad, el pensamiento filosófico se ocupa de los asuntos que importan. Para decidir qué deberíamos hacer, en vez de qué podemos hacer, tenemos que recurrir a la filosofía. Para aprender a vivir, para saber de justicia, de lenguaje, de estética, de realidad e irrealidad, para gozar, para amar, tenemos que recurrir a la filosofía.

En esta colección de 50 ensayos escritos de un modo tan accesible como lúcido, Ben Dupré nos presenta y nos explica todos estos conceptos que han merecido la atención de los pensadores, desde Antigua Grecia hasta el presente.

Ben Dupré

50 cosas que hay que saber sobre filosofía
50 cosas - 0

Introducción

Durante la mayor parte de su dilatada historia la filosofía ha contado con un número considerable de individuos peligrosos provistos de ideas peligrosas. A causa de sus ideas presuntamente subversivas, Descartes, Spinoza, Hume y Rousseau, por nombrar sólo a unos pocos autores, fueron amenazados con la excomunión, o forzados a aplazar la publicación de sus obras, o privados de las promociones profesionales, u obligados a exiliarse. Y al más notable de todos los filósofos, el ciudadano ateniense Sócrates, lo consideraron una influencia tan nociva que decidieron ejecutarlo. No hay muchos filósofos en la actualidad a los que se ejecute por sus ideas, lo cual es una lástima (en cuanto que indica hasta qué punto el sentido del peligro se ha ido desvaneciendo).

En la actualidad, la filosofía se considera la disciplina académica por antonomasia, con sus practicantes firmemente enclaustrados en sus torres de marfil, al margen de los problemas de la vida real. Pero la caricatura se encuentra lejos de la verdad en muchos sentidos. Los problemas de la filosofía son siempre profundos y a menudo difíciles, pero también *importan*. La ciencia, por ejemplo, tiene la capacidad de llenar el mercado con toda clase de golosinas, desde los niños de diseño hasta la comida modificada genéticamente, pero por desgracia no nos proporciona —y no puede hacerlo— el manual de instrucciones. Para decidir qué *deberíamos hacer*, en vez de qué podemos hacer, tenemos que recurrir a la filosofía. A veces, a los filósofos los mueve el placer de escucharse expnimiéndose el cerebro (e incluso puede resultar entretenido escucharles), pero por lo general aportan claridad y comprensión a asuntos que nos incumben a todos. Estos asuntos son precisamente los que este libro pretende reunir y explorar.

Tradicionalmente los autores suelen atribuir la mayor parte del mérito a los otros y se acusan de la mayor parte de los errores a sí mismos; tal vez sea una tradición, pero es un tanto ilógico (pues el mérito y los errores deberían darse la mano), y por lo tanto es una práctica difícilmente encomiable en un libro sobre filosofía. Así pues, con el mismo espíritu de P. G. Wodehouse al dedicar *The Heart of a Goof* a su hija, «sin cuya infatigable simpatía y estímulo hubiera terminado [el] libro en la mitad de tiempo», me complace por lo menos compartir el mérito con otros. En particular me alegra atribuírselo a mi jovial y trabajador editor, Keith Mansfield, por todas las cronologías y por las muchas citas bibliográficas que ha aportado. También me gustaría agradecer a mi editor en Quercus, Richard Milbank, su constante confianza y apoyo. Y mi mayor agradecimiento se lo debo a mi mujer, Geraldine, y a mis hijas, Sophie y Lydia, sin cuya infatigable simpatía…

01 El cerebro en una cubeta

Imaginad que un científico diabólico hubiera sometido a un experimento a un ser humano. Se habría extraído del cuerpo el cerebro de la persona y colocado en un recipiente con nutrientes que mantendría con vida el cerebro.
Las terminaciones nerviosas estarían conectadas a una computadora super científica capaz de provocar en la persona la ilusión de que todo es completamente normal. Parecería haber gente, objetos, el cielo, etc.; pero en realidad todo lo que la persona experimentaría sería el resultado de impulsos que van desde la computadora hasta las terminaciones nerviosas.

¿Se trata de una pesadilla de ciencia ficción? Tal vez, pero eso es exactamente lo que diríamos si fuéramos un cerebro metido en una cubeta. Si nuestro cerebro estuviera en un recipiente en vez de en el cráneo, cada una de nuestras experiencias sería exactamente igual que si hubiéramos vivido en un cuerpo real inmerso en el mundo real. El mundo circundante —esta silla, el libro que sostenéis con las manos, y las propias manos— forma parte de la ilusión, la poderosísima computadora del científico introduce en vuestros cerebros los pensamientos y las sensaciones.

Probablemente no creáis ser un cerebro flotando en una cubeta. Es posible que la mayoría de los filósofos no crean ser cerebros en cubetas. Pero no se trata de que lo creamos sino tan sólo de admitir que no es posible tener la certeza de que no lo somos. El problema es que, si realmente somos un cerebro en una cubeta (simplemente no podemos descartar la posibilidad), todas las cosas que creemos conocer del mundo serían falsas. La mera posibilidad parece minar nuestras pretensiones de conocimiento acerca del mundo exterior. ¿Existe alguna forma de escapar de la cubeta?

Cronología

c. 375 a. C.	1637	1644
La caverna platónica	El problema del cuerpo y la mente	*Cogito ergo sum*

Los orígenes de la cubeta El clásico y elocuente relato moderno del « cerebro en una cubeta» lo urdió el filósofo norteamericano Hilary Putnam en su libro *Razón, verdad e historia* (1981), pero el germen de la idea se remonta mucho más atrás. El experimento mental de Putnam actualiza una historia de terror del siglo XVII (el genio maligno —*malin génie*—, convocado por el filósofo francés René Descartes en sus *Meditaciones* de 1641). El propósito de Descartes consistía en edificar el conocimiento humano sobre fundamentos inquebrantables, para lo cual adoptó la « duda metódica» (desechaba cualquier creencia susceptible del

menor grado de incertidumbre). Tras señalar el carácter engañoso de nuestros sentidos y la confusión propia de los sueños, Descartes llevó su « duda» hasta el límite:

« Debo suponer… que algún genio maligno inmensamente poderoso y astuto ha dedicado todas sus energías a engañarme. Debo pensar que el cielo, el aire, la tierra, los colores, las formas, los sonidos y todas las cosas externas son meras ilusiones oníricas que este genio ha inventado para cautivar mi juicio» . Entre los escombros de sus antiguas creencias y opiniones, Descartes vislumbra un solo punto de certeza —el *cogito*— en el que fundar de un modo (aparentemente) seguro la reconstrucción que se ha propuesto como tarea (véase página 20).

Desgraciadamente para Putnam y Descartes, aunque ambos están haciendo de abogado del diablo —al adoptar posiciones escépticas para frustrar el escepticismo—, a algunos filósofos les ha impresionado más su habilidad para plantear el atolladero del escepticismo que sus posteriores tentativas para salir de él. Apelando a su propia teoría causal del significado, Putnam intenta mostrar que la escena del cerebro en una cubeta es incoherente, pero a lo sumo parece conseguir mostrar que de hecho un cerebro en una cubeta no podría expresar el pensamiento de ser un cerebro en una cubeta. Efectivamente, demuestra que el estado de ser un cerebro envasado es invisible e indescifrable para el espíritu, pero no está claro que esta victoria semántica (si lo es) consiga resolver el problema relativo al conocimiento.

En la cultura popular

Ideas como la del cerebro en una cubeta han resultado tan estimulantes humanos los mantienen en el interior y sugestivas intelectualmente que han dado lugar a numerosas manifestaciones populares. Una de las más exitosas fue la película *Matrix* en 1999, en la que un *hacker* llamado Neo (interpretado por Keanu Reeves) descubre que el mundo americano de 1999 es, de hecho, una simulación virtual creada por una ciber inteligencia maligna, y que a él y otros seres humanos los mantienen en el interior de recipientes llenos de un fluido, conectados a una inmensa computadora. La película presenta una elaboración dramática del cerebro en una cubeta, en la que se reproducen los principales elementos de la situación. El éxito y el impacto de *Matrix* advierte de la fuerza que poseen los argumentos del escepticismo radical.

El escepticismo El término « escéptico» se aplica comúnmente a las personas con tendencia a dudar de las creencias aceptadas, o habituadas a desconfiar de la gente o de las ideas en general. En este sentido, el escepticismo puede caracterizarse como una tendencia saludable o un ejercicio propio de mentes abiertas que consiste en someter a prueba y demostrar las creencias comúnmente aceptadas. Un estado mental semejante suele ser una salvaguarda útil contra la credulidad, pero a veces también puede desembocar en la tendencia a dudar de todo, con independencia de las razones para hacerlo. Pero sea bueno o malo, el escepticismo en este sentido común es bastante distinto al escepticismo en sentido filosófico.

El argumento de la estimulación

La gente corriente podría estar tentada de desdeñar las terroríficas conclusiones escépticas, pero no deberíamos apresurarnos. Por lo demás, un ingenioso argumento inventado recientemente por el filósofo Nick Bostrom sugiere que es muy probable que ya estemos viviendo en una simulación informática. Imaginad que…

Es probable que en el futuro nuestra civilización alcance tal nivel de desarrollo tecnológico que puedan crearse simulaciones informáticas muy sofisticadas de mentes humanas y de mundos donde habiten esas mentes. Los recursos que precisaría el mantenimiento de esos mundos serían podría relativamente insignificantes —un simple ordenador portátil del futuro podría ser el hogar de cientos o de miles de mentes simuladas — de modo que muy probablemente el número de mentes simuladas superaría con mucho el de las biológicas. Las experiencias de las mentes biológicas y las de las simuladas serían indiscernibles y, como es lógico, cada una de ellas pensaría que no es simulada, pero las segundas (que constituirían la amplia mayoría de las mentes) estarían equivocadas. Naturalmente ponemos a prueba este argumento como hipótesis de futuro, pero ¿quién puede afirmar con seguridad que este « futuro» no ha llegado (que tal logro informático no se ha alcanzado ya, y que no existen mentes simuladas)? Es evidente que suponemos no ser mentes simuladas informáticamente e inmersas en un mundo simulado, pero esto podría tratarse como tributo a la calidad del programa. De acuerdo con la lógica del argumento de Bostrom ¡es muy probable que nuestra suposición sea errónea!

El escéptico filosófico no pretende que no sepamos nada (en buena medida porque pretenderlo resultaría obviamente contradictorio: no podemos saber que no sabemos nada). La posición escéptica consiste más bien en cuestionar nuestro derecho a pretender algún conocimiento. Creemos saber muchas cosas, pero ¿cómo podemos defender esa pretensión? ¿Qué solidez podemos ofrecer para justificar cualquier afirmación concreta relativa al conocimiento? Nuestro supuesto conocimiento del mundo se basa en percepciones que nos proporcionan nuestros sentidos, por lo general mediadas por nuestro uso de razón. Pero ¿acaso esas percepciones no se encuentran sometidas en ocasiones al error? ¿Podemos estar completamente seguros de que no estamos sumidos en una alucinación o en un sueño, o de que nuestra memoria no nos tiende trampas? Si la experiencia del sueño es indiscernible de la experiencia de la vigilia, nunca podremos tener la certeza de que algo que pensamos que es, sea de hecho (ni de que lo que consideramos cierto lo sea). Estas inquietudes, llevadas al extremo, desembocan en los genios malignos y en los cerebros en cubetas…

«**La computadora es tan inteligente que a la víctima incluso podría parecerle que se encuentra sentada y leyendo estas mismas palabras sobre el supuesto, divertido aunque más bien absurdo, de un científico maligno que saca los cerebros de los cuerpos de la gente para ponerlos en una cubeta llena de nutrientes.**»

Hilary Putnam, 1981

La epistemología, el ámbito de la filosofía consagrado al conocimiento, determina qué sabemos y cómo lo sabemos, e identifica en qué condiciones algo debe ser conocido para ser considerado conocimiento. Así entendida, puede concebirse como una respuesta al desafío del escepticismo; y su historia como las distintas tentativas de derrotar al escepticismo. A muchos autores les parece que ha habido pocos filósofos que hayan conseguido vencer al escepticismo mejor que Descartes. La posibilidad de que en el fondo no exista una vía de salida segura de la cubeta sigue proyectando una larga sombra sobre la filosofía.

La idea en síntesis: ¿somos cerebros envasados?

02 La caverna de Platón

«Imagina que toda tu vida has sido prisionero en una caverna. Tienes las manos y los pies encadenados, y la cabeza sujeta de modo que sólo puedes ver la pared que queda enfrente. Detrás de ti hay una llama, y entre tú y el fuego una pasarela por la que tus captores desplazan estatuas y todo tipo de objetos. Las sombras que proyectan en la pared estos objetos son lo único que tú y tus compañeros de cautiverio habéis visto siempre, lo único de lo que habéis hablado y en lo que habéis pensado.»

De entre las muchas imágenes y analogías que utilizó el filósofo griego Platón, posiblemente el mito de la caverna sea la más célebre: aparece en el libro VII de *La República*, una obra monumental en la que investiga la forma del Estado ideal y su gobernante ideal, el filósofo rey. La justificación de Platón para dar las riendas del gobierno a los filósofos se encuentra en un pormenorizado estudio en torno de la verdad y el conocimiento, y en este contexto es donde interviene la alegoría de la caverna.

La concepción platónica del conocimiento y de sus objetos es compleja y se compone de muchos estratos, tal como evidencia el modo en que prosigue la parábola de la caverna.

Cronología

c. 375 a. C.	1644
La caverna de Platón	*Cogito ergo sum*

« Ahora supón que te liberan de las cadenas y puedes andar por la cueva libremente. Aunque al principio el fuego te deslumbra, de forma progresiva vas reconociendo mejor la situación de la cueva y entendiendo el origen de las sombras que habías tomado por reales. Y finalmente se te permitirá salir de la caverna y asomarte al soleado mundo exterior, donde verás toda la realidad iluminada por el cuerpo más brillante que hay en los cielos, el Sol.»

La interpretación de la caverna Se ha discutido mucho sobre la interpretación detallada de la caverna de Platón, pero el significado general está bastante claro. La caverna representa « el mundo de las apariencias» (el mundo visible de nuestra experiencia cotidiana, donde todo es imperfecto y constantemente cambiante). Los cautivos encadenados (que simbolizan la gente corriente) viven en un mundo de conjeturas e ilusión,

«¡Contemplad! Son seres humanos que viven en una guarida subterránea … Como nosotros … Sólo ven sus propias sombras, o las sombras de los demás, que el fuego proyecta en la pared de enfrente de la caverna.»

mientras que el antiguo prisionero, libre para deambular por el interior de la caverna, obtiene la visión más adecuada posible de la realidad en el seno del cambiante mundo de la percepción y la experiencia. En cambio, el exterior de la caverna representa «el mundo de las ideas» (el mundo inteligible de la verdad poblado de los objetos del conocimiento, que es perfecto, eterno e inmutable).

La teoría de las Ideas Para Platón, lo que conocemos no sólo debe ser verdadero sino también perfecto e inmutable. Sin embargo nada en el mundo empírico (representado por la vida en el interior de la caverna) se adecúa a esta descripción: una persona alta es bajita al lado de un árbol; una manzana que parece roja al mediodía se ve negra al anochecer; y así ocurre con todo. Puesto que nada en el mundo empírico es un objeto de conocimiento, Platón propone que debe de existir otro reino (el mundo en el exterior de la caverna) de entidades perfectas e inmutables a las que llama «Ideas» (o Formas). Así, por ejemplo, sólo gracias a la virtud de copiar o imitar la Idea de la Justicia es justa toda acción concreta que sea justa. Como sugiere el mito de la caverna, entre las ideas existe una jerarquía, y gobernándolas a todas se encuentra la Idea de Dios (que representa el Sol), que brinda a las otras su sentido último e incluso subyace a su existencia.

El amor platónico

La Idea con que se asocia más comúnmente a Platón en la actualidad —el llamado amor platónico— surge de un modo natural del marcado contraste establecido en el mito de la caverna entre el mundo del intelecto y el mundo de los sentidos. La clásica afirmación de la idea de que el tipo de amor más perfecto se expresa no física sino intelectualmente se encuentra en otro célebre diálogo. *El Banquete*.

1690	1981
El velo de la percepción	El cerebro en una cubeta

El problema de los universales La teoría de las ideas de Platón y la base metafísica que implica puede resultarnos exótica y excesivamente complicada, pero el problema que intenta resolver —el llamado «problema de los universales»— ha sido, con variaciones, uno de los principales problemas de la filosofía desde siempre. En la Edad Media los bandos de la batalla se dividían entre los realistas (o platónicos) de un lado, quienes creían que universales como la condición de «ser de color rojo» y de «ser alto» existían con independencia

de los objetos particulares rojos o altos; y los nominalistas por otro lado, quienes sostenían que tales cosas eran meros nombres o etiquetas que asociamos a objetos para establecer similitudes particulares entre ellos.

En la cultura popular

Existe un claro eco del platónico mito de la caverna en la obra de C. S. Lewis, el autor de las siete novelas fantásticas que forman *Las crónicas de Narnia*. Al final del último libro. *La última batalla*, los niños que protagonizan la historia atestiguan la destrucción de Narnia y escapan a las tierras de Aslan, un país maravilloso que reúne lo mejor de la antigua Narnia y de la Inglaterra que recuerdan. Los niños terminan descubriendo que están muertos y que han abandonado « la tierra de las sombras» , que no era más que una pálida imitación del mundo eterno e inmutable en el que ahora moran. A pesar del evidente mensaje cristiano que encierra la historia, la influencia de Platón es clara: se trata de uno de los innumerables ejemplos de la enorme (y a menudo inesperada) influencia del filósofo griego en la cultura, la religión y el arte occidentales.

La misma distinción fundamental, habitualmente expresada en términos de realismo y antirrealismo, sigue teniendo eco en muchos campos de la filosofía moderna. Así, una posición realista sostiene que existen entidades « fuera» del mundo —cosas físicas o actos éticos o propiedades matemáticas— que existen independientemente de nuestro conocimiento o de nuestra experiencia. A este punto de vista se opone el de otros filósofos, conocidos como antirrealistas, que plantean propuestas en las que existe una relación necesaria e interna entre las cosas que conocemos y nuestro conocimiento de las mismas. Los términos fundamentales de todos estos debates los estableció hace 2 000 años Platón, uno de los primeros (y más concienzudos) realistas filosóficos.

En defensa de Sócrates Con su mito de la caverna Platón se propone hacer algo más que sacar a la luz sus puntos de vista particulares sobre la realidad y sobre nuestro conocimiento acerca de ella. Algo que queda claro al final de su mito. Una vez ha ascendido al mundo exterior y reconocido la naturaleza de la verdad y la realidad últimas, el prisionero liberado está ansioso por volver a la caverna y sacar del error a sus antiguos compañeros ignorantes. Pero como ahora está acostumbrado al brillo de la luz del mundo exterior, al principio va dando traspiés, inmerso en la oscuridad de la caverna, y quienes siguen allí cautivos lo

consideran un loco. Piensan que su viaje lo ha trastocado; se niegan a escucharle e incluso podrían terminar matándole si persiste. En este pasaje Platón alude al tópico de la complicada situación del filósofo —la burla y el rechazo— cuando intenta ilustrar a la gente corriente y orientarla hacia el conocimiento y la sabiduría. Y en particular se refiere al destino de su maestro, Sócrates (su portavoz en la *República* y en muchos otros de sus diálogos), quien rechazó a lo largo de toda su vida moderar sus enseñanzas filosóficas, y al que la ciudad de Atenas terminó ejecutando en el año 399 a. C.

La idea en síntesis: el conocimiento en este mundo no es más que una sombra

03 El velo de la percepción

¿Cómo vemos (y oímos y olemos) el mundo?

La mayoría de nosotros damos por descontado que los objetos físicos que nos rodean son más o menos como los percibimos, pero esta noción del sentido común plantea problemas que han llevado a algunos filósofos a preguntarse si es cierto que observamos el mundo exterior directamente. Desde su punto de vista sólo tenemos acceso a «ideas» o «impresiones» interiores o (en términos modernos) a «datos de los sentidos».

El filósofo inglés del siglo XVII John Locke acuñó una célebre metáfora para elucidar este asunto. Sugirió que el conocimiento humano es una especie de «armario completamente oscuro, con unas pocas aberturas tan sólo, por las que entran apariencias externas visibles, o ideas de cosas del exterior».

Pero la concepción de Locke plantea un problema grave. Podemos suponer que las ideas que entran en el armario son representaciones más o menos fieles de las cosas que hay afuera, pero en última instancia el que esas representaciones interiores se correspondan fielmente a los objetos externos (o a cualquier cosa, al fin y al cabo) es una inferencia. Nuestras ideas, que es a lo único a lo que tenemos acceso, forman un impenetrable «velo de la percepción» entre nosotros y el mundo exterior.

En su *Ensayo sobre el entendimiento humano* de 1690, Locke dio una de las explicaciones más exhaustivas de lo que se conoce como modelos «representacionales» de la percepción. Cualquier modelo que, como éste, incorpore ideas mediadoras o datos de los sentidos abre una brecha entre nosotros y el mundo exterior, y es en esta brecha en la que echa sus raíces el escepticismo sobre nuestras pretensiones de conocimiento. Sólo restableciendo un vínculo directo entre el observador y los objetos externos puede rasgarse el velo y derrotarse al escepticismo. Pero considerando todos los problemas que provoca este modelo, ¿por qué adoptarlo de entrada?

Cronología

c. 375 a. C.	1644
La caverna de Platón	*Cogito ergo sum*

El teatro cartesiano

En la modernidad se llama «realismo representativo» al modelo de percepción de Locke, para distinguirlo del realismo «ingenuo» (o del «sentido común») que adhiere cualquiera de nosotros la mayor parte

del tiempo (e incluso los filósofos cuando no ejercen). Las dos posiciones son realistas en la medida en que asumen la existencia del mundo exterior con independencia de nosotros, pero sólo en la versión ingenua el color rojo es considerado como una mera propiedad del tomate mismo. Aunque Locke pesar sea posiblemente el primero que dio forma a la teoría, el modelo representacional de la percepción no se inició con él. A veces se califica despectivamente de « teatro cartesiano» , porque, en efecto, para Descartes la mente es un escenario en el que las ideas (las percepciones) son observadas por un observador interior (el alma inmaterial). El hecho de que este observador interior, u « homúnculo» , requiera también un observador interior para sí mismo (y así hasta el infinito) es sólo una de las objeciones que cabe plantear a esta teoría. Sin embargo, a pesar de estas objeciones, el modelo sigue siendo muy influyente.

Cualidades primarias y secundarias La falta de fiabilidad de nuestras percepciones constituye una parte importante del arsenal del escepticismo para atacar nuestras pretensiones de conocimiento. El escéptico se atiene al hecho de que un tomate pueda verse rojo o negro dependiendo de las variaciones de luz para cuestionar en general nuestros sentidos como una vía fiable de conocimiento. Locke confía en que un modelo perceptual en el que las ideas interiores y los objetos exteriores se mantuvieran separados desarmaría al escéptico. Su argumento dependía fundamentalmente de una distinción adicional entre las cualidades primarias y secundarias.

1690	1981
El velo de la percepción	El cerebro en una cubeta

«**El conocimiento humano … no puede ir más allá de la experiencia.**»

John Locke, 1690

El color rojo del tomate no es una propiedad del tomate en sí misma sino el producto de la interacción de varios factores, incluidos determinados atributos físicos del tomate, como su textura y la estructura de su superficie; las peculiaridades de nuestro sistema sensitivo; y las condiciones ambientales imperantes en el momento de la observación. Tales propiedades (o mejor dicho, no propiedades) no pertenecen al tomate como tal y se dice que son « cualidades secundarias» .

Al mismo tiempo, el tomate tiene algunas propiedades verdaderas, como su forma y su tamaño, que no dependen de las condiciones en las que se observa, ni siquiera de la existencia de un observador. Éstas son sus « cualidades primarias» ,

que explican y dan lugar a nuestra experiencia de las cualidades secundarias. A diferencia de lo que ocurre con nuestras ideas de las cualidades secundarias, las de las cualidades primarias (pensaba Locke) se parecen mucho a los objetos físicos en sí mismos y nos proporcionan su conocimiento. Por esta razón son las cualidades primarias las que más preocupan a los científicos, y las que resultan más cruciales para combatir el desafío del escepticismo, pues son nuestras ideas de las cualidades primarias las que constituyen una prueba contra las dudas escépticas.

«Así es como lo refuto»

En la actualidad la teoría inmaterialista de Berkeley se considera un *tour de force* de un gran virtuosismo, aunque también un tanto excéntrico. Aun así, irónicamente Berkeley se veía *a sí mismo* como el campeón del sentido común. Tras exponer con gran habilidad los inconvenientes de la concepción mecanicista del mundo en Locke, propuso una solución que le parecía obvia y que resolvía todos esos inconvenientes de un plumazo, disipando tanto la inquietud de los escépticos como la de los ateos. Pero lo que más hubiera mortificado a Berkeley hubiera sido saber que su lugar en el imaginario popular actual se limita al célebre pensamiento con el que Samuel Johnson refutó firmemente el inmaterialismo, que se encuentra en *Vida de Samuel Johnson* de Boswell: « Golpeando con todas sus fuerzas una gran piedra con el pie [exclamó]: "¡Así es como lo refuto!"» .

Encerrados en el armario de Locke Uno de los primeros críticos de Locke fue su contemporáneo irlandés, George Berkeley. Éste asumía el modelo representacional de la percepción en el que los objetos inmediatos de la percepción eran las ideas, pero al mismo tiempo sostenía que, lejos de contradecir a los escépticos, la concepción de Locke corría el riesgo de lograr que cualquiera se rindiera a ellos. Puesto que Locke estaba enclaustrado en su armario, nunca conseguiría comprobar si las supuestas « similitudes o las ideas de las cosas externas» se asemejaban realmente a las cosas externas. Nunca podría rasgar el velo y mirar del otro lado, de modo que se encontraba atrapado en un mundo de representaciones, y el escepticismo estaba servido.

«Una opinión extrañamente difundida entre los hombres es que las casas, las montañas, los ríos, y en una palabra todos los objetos sensibles, tienen una existencia natural o real, distinta de su ser percibido.»

George Berkeley, 1710

Berkeley, que había expuesto los peligros de la posición de Locke de un modo lúcido, llegó a una extraordinaria conclusión. En vez de rasgar el velo en una tentativa de restablecer el vínculo entre nosotros y el mundo exterior, concluyó ¡que no había nada tras el velo! Para Berkeley, la realidad consiste en las propias «ideas» y las propias sensaciones. Así, está claro, ya estamos completa y adecuadamente vinculados, de modo que se evitan los peligros del escepticismo, aunque al precio… ¡de negar la existencia del mundo físico exterior!

De acuerdo con la teoría idealista (o inmaterialista) de Berkeley, «existir es ser percibido» (*ese est percipi*). Pero entonces ¿dejan de existir las cosas cuando dejamos de mirarlas? Berkeley acepta esta consecuencia, pues la solución está muy a mano: Dios. Todo lo que existe en el universo es concebido continuamente en la mente de Dios, de manera que la existencia y la continuidad del mundo (inmaterial) están garantizadas.

La idea en síntesis: ¿que hay detrás del velo?

04 Cogito ergo sum

Desprovisto de cualquier creencia de la que se le ocurriera que cabía dudar, a la deriva en un mar que parece de una incertidumbre insondable, Descartes trata de encontrar desesperadamente algún punto de apoyo, algún fundamento sólido sobre el que reconstruir el edificio del conocimiento humano…

« Advertí que cuando intentaba pensar que todo era falso, era necesario que yo, que estaba pensándolo, fuera algo. Y al observar que esta verdad, "Pienso, luego existo [*cogito ergo sum*]" era tan firme y segura que ni siquiera el supuesto más extravagante de los escépticos era capaz de amenazarla, decidí que podía aceptarla sin reservas como el primer principio de la filosofía que estaba buscando.»

Así llegó este francés llamado René Descartes a pensar el pensamiento sin duda más célebre y más influyente de la historia de la filosofía occidental.

«**Je pense, donc je suis.**»

René Descartes, 1637

La duda metódica El propio Descartes se encontraba a la vanguardia de la revolución científica en la que estaba inmersa Europa en el siglo XVII, y su ambicioso plan consistía en dejar atrás los viejos dogmas del mundo medieval y « establecer las ciencias» sobre los fundamentos más sólidos. Con este propósito adoptó su rigurosa « duda metódica» . No se contentó con separar la extraña manzana podrida (en sus propias palabras), sino que vació completamente el tonel, desechando cualquier creencia que fuera susceptible del menor grado de duda. Y en una última vuelta de tuerca imaginó a un genio maligno dedicado exclusivamente a confundirlo, de modo que ni siquiera las verdades aparentemente evidentes de la geometría y de las matemáticas podían seguir siendo ciertas.

Cronología

1637	1644
El problema del cuerpo y la mente	*Cogito ergo sum*

El cogito en las obras de Descartes

La conocida expresión en latín *cogito ergo sum* se encuentra en los *Principios de filosofía* (1644) de Descartes, pero en su *Discurso del método* (1637) aparece la versión francesa (*je pense, donc je suis*) y en su obra más importante, las *Meditaciones*, no aparece en ninguna de estas formas canónicas.

Llegado a este punto —despojado de todo, incluso de su propio cuerpo y de los sentidos, de los demás, y de todo el mundo externo—, Descartes encuentra la salvación en el *cogito*. Por más engañado que pueda estar, por más que se empeñe el genio en confundirlo en cada ocasión, tiene que haber alguien o algo a quien confundir, algo o alguien a quien engañar. Incluso si está equivocado sobre cualquier otra cosa, no puede dudar de que él está ahí, en ese momento, para albergar el pensamiento de que podría estar equivocado. El genio «nunca podría hacer que yo fuera nada desde el momento en que pienso que soy algo … *Soy, existo*, es necesariamente verdadero en todas las ocasiones en que me lo planteo o lo concibo en mi mente».

Los límites del *cogito* Una de las primeras críticas a Descartes, retomada por muchos otros autores desde entonces, es que infiere demasiadas cosas del *cogito*: sólo puede concluir legítimamente que hay pensamiento, pero no que sea él quien está pensando. Pero incluso si concedemos que los pensamientos efectivamente presuponen un individuo que los piense, debe reconocerse que lo que establece el hallazgo de Descartes es muy limitado. En primer lugar, el *cogito* es esencialmente relativo a la primera persona (mi *cogito* sólo es aplicable a mí y el tuyo a ti: es perfectamente posible que el genio tenga poderes para hacerme creer que *tú* piensas, y por lo tanto que existes). En segundo lugar, el *cogito* se da esencialmente en tiempo presente: en consecuencia, es perfectamente posible que yo deje de existir cuando no pienso. En tercer lugar, el «yo» cuya existencia se establece es extremadamente tenue e inaprensible: podría no tener ninguno de los atributos biográficos o de otro tipo que me hacen creer que yo soy yo y, en efecto, podría seguir perfectamente en las garras del genio maligno.

1690	1981
El velo de la percepción	El cerebro en una cubeta

Orígenes del «cogito»

Cogito ergo sum es posiblemente el dicho filosófico más célebre, pero sus orígenes precisos no están demasiado claros. Aunque se encuentra inextricablemente unido a Descartes, la idea que subyace al *cogito* se remonta mucho más atrás. A principios del siglo V d. C., por ejemplo, san Agustín escribió que podemos dudar de todo menos de la duda de nuestra propia alma, y la idea no la inventó él.

En suma, el «yo» del *cogito* es un mero instante de autoconciencia, un mero atisbo separado de todo lo demás, incluido su propio pasado. ¿Qué puede establecer Descartes sobre una base tan precaria?

La reconstrucción del conocimiento Descartes tal vez despejara el terreno para cavar los cimientos, pero ¿acaso le quedaba algún material con el que erigir el nuevo edificio? En apariencia, había colocado el listón demasiado alto (nada menos que a la altura de una certeza a prueba de genios malignos). Y así, el viaje de regreso resulta sorprendentemente (tal vez alarmantemente) rápido. Los principales pilares de la teoría cartesiana del conocimiento son dos. En primer lugar, señala que el rasgo distintivo del *cogito* es la claridad con que se nos presenta y, de ahí, concluye que existe una regla general: «todas las cosas que concebimos con claridad y distinción son verdaderas». ¿Y cómo podemos estar seguros de esto? Porque la idea más clara y distinta de todas es nuestra idea de un Dios perfecto, todopoderoso y omnisciente.

Dios es la fuente de todas nuestras ideas y, puesto que es bueno, no puede engañarnos; el uso de nuestra capacidad de observación y racional (que también proceden de Dios) nos conduce hacia la verdad, no hacia lo falso. Merced a la intervención de Dios, los mares de la duda se retiran rápidamente: el mundo queda restaurado y la tarea de reconstrucción del conocimiento puede iniciarse sobre una base sólida, científica.

«… recurrir a la veracidad del Ser supremo para probar la veracidad de nuestros sentidos es, sin duda, optar por un camino inesperado.»

David Hume, 1748

Dudas persistentes El intento cartesiano de salir del pozo escéptico que él mismo ha cavado está lejos de convencer a todo el mundo. Las críticas se han centrado en el lamentablemente célebre «círculo cartesiano»: el supuesto recurso a las ideas claras y distintas para probar la existencia de Dios, cuya divinidad acaba garantizando el recurso a nuestras ideas claras y distintas. Por más poderoso que sea el argumento (y no está nada claro que Descartes cayese, de hecho, en una trampa tan evidente), resulta difícil compartir su confianza en haber logrado exorcizar al genio maligno. Descartes no puede (y no lo hace) negar el hecho de que el engaño tiene lugar; y si nos atenemos a su regla general, ella implica que podemos equivocamos en ocasiones al pensar que tenemos una idea clara y distinta de algo. Pero, naturalmente, no podemos saber que estamos incurriendo en tal error, y si no somos capaces de identificar estos casos, la puerta del escepticismo queda una vez más completamente abierta.

A Descartes se le ha considerado el padre de la filosofía moderna. Merece con creces el título, pero seguramente no por las razones que a él le hubiera gustado. Su propósito era disipar las dudas escépticas de una vez por todas, para podernos dedicar tranquilamente a la búsqueda racional del conocimiento, pero al final tuvo más éxito planteando estas dudas que sofocándolas. Las sucesivas generaciones de filósofos quedaron atrapadas en el problema del escepticismo, que ocupa la primera posición, o está muy cerca, en el orden del día filosófico desde que Descartes lo planteó.

La idea en síntesis: pienso, luego existo

05 Razón y experiencia

¿Cómo llegamos a conocer las cosas? ¿Adquirimos conocimientos principalmente a través del uso de razón? ¿O la experiencia recabada a través de nuestros sentidos juega el papel más significativo en la manera en que conocemos el mundo? La oposición entre razón y experiencia como principio fundacional del conocimiento ha teñido la mayor parte de la historia de la filosofía occidental. Constituye específicamente la manzana de la discordia entre dos tendencias filosóficas muy influyentes: el racionalismo y el empirismo.

Tres distinciones clave Para entender lo que está en juego en las teorías racionalistas y empiristas del conocimiento, resulta útil tener en cuenta tres distinciones clave que utilizan los filósofos para elucidar sus diferencias.

a priori versus a posteriori

Algo es cognoscible a priori si puede ser conocido sin remisión a la experiencia, es decir, sin ninguna investigación empírica de cómo ocurren realmente las cosas en el mundo: « 2 + 2 = 4» es algo que conocemos a priori: no es preciso observar el mundo para establecer su verdad. Por el contrario, cuando se requiere tal investigación nos encontramos ante algo únicamente cognoscible a posteriori; así « el carbón es negro» , aun siendo cierto, es una verdad a posteriori: para verificarla es necesario mirar un pedazo de carbón.

analítico versus sintético

Una proposición es analítica si no proporciona ninguna otra información que la que ya contienen los significados de sus términos. La verdad de la afirmación « Todas las solteras no están casadas» es evidente tan sólo en virtud de la comprensión del significado y de la relación entre las palabras usadas. En cambio, la afirmación « Todas las solteras son desdichadas» es sintética: reúne (sintetiza) distintos conceptos y así proporciona información significativa (o, en este caso, desinformación). Para establecer si es verdadera o no sería necesario indagar el estado mental de cada mujer soltera.

Cronología

c. 350 a. C.	1670
Formas del argumento	Fe y razón

necesario versus contingente

Una verdad necesaria es la que no puede ser de otro modo: debe ser verdadera en todas las circunstancias o en todos los mundos posibles. Una verdad contingente es verdadera pero podría no haberlo sido si las cosas en el mundo

hubieran sido distintas. Por ejemplo, la afirmación «La mayoría de los chicos son gamberros» es contingente: puede ser cierta o no serlo, dependiendo de cómo se comporten la mayoría de los jóvenes. En cambio, si es cierto que todos los chicos son gamberros y Ludwig es un chico, es necesariamente verdadero (en este caso es una cuestión de lógica) que Ludwig es un gamberro.

Las preocupaciones kantianas

La distinción entre analítico/sintético tiene sus orígenes en la obra del filósofo alemán Immanuel Kant. Uno de sus principales objetivos en la *Crítica de la razón pura* es demostrar que existen ciertos conceptos o categorías del pensamiento, como el de sustancia y el de causa, que no podemos aprender del mundo pero de los que precisamos para que tenga sentido. El principal tema en Kant es la naturaleza y la justificación de estos conceptos, y del conocimiento sintético a priori a que dan lugar.

1739	1781	1963
Ciencia y pseudociencia	Razón y experiencia El imperativo categórico	La teoría tripartita del conocimiento

«El pilar en que se apoyan las matemáticas es puramente metafísico.»

Thomas de Quincey, 1830

Parece existir un evidente paralelismo entre estas distinciones: así, a primera vista, cuando un juicio analítico es verdadero, lo es necesariamente y es conocido a priori; y cuando un juicio sintético es verdadero, lo es de un modo contingente y es conocido a posteriori. Sin embargo, de hecho, las cosas no son tan claras, y la principal diferencia entre empiristas y racionalistas puede establecerse a partir del modo en que disponen estos términos. Así, la tarea de los racionalistas consiste en mostrar que existen juicios sintéticos a priori, es decir, que es posible descubrir racionalmente, sin recurso a la experiencia, hechos significativos o dotados de sentido sobre el mundo. Por el contrario, el propósito de los empiristas suele ser mostrar que los hechos a priori, como los de las matemáticas, son analíticos (véase el cuadro).

Alternativas a la fundamentación Los racionalistas y los empiristas pueden diferir en muchos puntos, pero están de acuerdo al menos en que existe una *base* (la razón o la experiencia) en la que se funda nuestro conocimiento. Así, por ejemplo, el filósofo escocés del siglo XVIII David Hume criticaba la quimérica busca cartesiana de una firme certeza racional que le permitiera afirmar todo nuestro conocimiento, incluida la veracidad de nuestros sentidos. Pero Hume no

niega que exista *alguna* base, sino simplemente que tal fundamentación pueda excluir nuestra experiencia corriente y los sistemas de creencias naturales.

El campo de batalla de las matemáticas

En el conflicto entre el empirismo y el racionalismo, las matemáticas se han convertido en el campo de batalla donde han tenido lugar las luchas más encarnizadas. Para los racionalistas, las matemáticas siempre han ofrecido un paradigma del conocimiento, al presentar un reino de objetos abstractos en el que los descubrimientos pueden tener lugar mediante el exclusivo ejercicio de la razón. Un empirista no puede permitir no cuestionar esta posición, de modo que está obligado o bien a negar que los hechos matemáticos puedan conocerse de este modo, o bien a mostrar que tales hechos son esencialmente analíticos y triviales. La última de las opciones normalmente consiste en argumentar que los presuntos hechos abstractos de las matemáticas son, en realidad, constructos humanos y que el pensamiento matemático es, en el fondo, el producto de la convención: en última instancia descansa en el consenso, no en el descubrimiento; se trata de pruebas no de la verdad.

De modo que tanto el racionalismo como el empirismo tienen una vocación fundamentadora común, pero hay otras aproximaciones que prescinden de este supuesto básico. Una alternativa que ha tenido un gran influjo es el coherentismo, en el que el conocimiento es considerado como una trama de creencias, cuyos hilos forman un tejido o una estructura coherente. Sin embargo, sigue siendo una estructura sin un fundamento único, de ahí el principio del coherentismo: «todo argumento requiere premisas, pero no existe nada que sea la premisa de todos los argumentos».

Rivalidades europeas

Desde el punto de vista histórico, los empiristas ingleses de los siglos XVII y XVIII —Locke, Berkeley y Hume— suelen agruparse contra sus «rivales» continentales, los racionalistas Descartes, Leibniz y Spinoza. Pero como es habitual, tales categorizaciones simplistas oscurecen buena parte de los matices. El arquetipo de un bando. Descartes, suele manifestar su simpatía por las investigaciones empíricas, mientras que Locke, el arquetipo del otro bando, se muestra a veces dispuesto a

conceder el espacio que los racionalistas quisieran atribuir a alguna forma de agudeza racional o intuición.

La idea en síntesis: ¿cómo conocemos?

06 La teoría tripartita del conocimiento

«Oh, vaya, qué giro inesperado», pensó Don mientras veía la detestable figura desplomarse contra la farola, y los rasgos perfectamente familiares de su rostro británico claramente visibles bajo la luz amarillenta. «Debería haber previsto que esta escoria aparecería. Bueno, ya lo sé… ¿A qué esperas, Eric? ¿No eras tan duro…?» Don estaba completamente concentrado en la figura que tenía delante, y no oyó los pasos que se le acercaban por detrás. Ni sintió nada cuando Eric le propinó el golpe fatal en la nuca…

¿Sabía Don realmente que su asesino Eric se encontraba en el callejón aquella noche? Don creía, en efecto, que se encontraba allí, y su creencia terminó siendo cierta. Disponía de todas las razones para creerlo: pero no tenía la menor idea de que Eric tuviera un hermano gemelo idéntico llamado Alee, y vio a un hombre perfectamente idéntico a Eric.

La definición platónica del conocimiento Nuestra intuición nos indica que Don no sabía que Eric se encontraba en el callejón: pese a que Eric estaba efectivamente allí, Don creía que lo tenía delante y, en principio, su creencia estaba plenamente justificada. Pero al decir esto contrariamos una de las definiciones de la historia de la filosofía más consagradas.

En su diálogo *Teeteto*, Platón lleva a cabo una exhaustiva investigación sobre el concepto de conocimiento. La conclusión a la que llega es que el conocimiento es «una creencia verdadera con un *logos*» (es decir, con una «explicación racional» de por qué la creencia es verdadera), o simplemente «una creencia verdadera justificada». Esta teoría del conocimiento, llamada tripartita, puede expresarse formalmente del siguiente modo:

Cronología

c. 350 a. C.	1781
Formas del argumento	Razón y experiencia

Un individuo S conoce la proposición *P* si y sólo si:

1. *P* es verdadera.
2. S cree que *P*.
3. La creencia de S en *P* está justificada.

De acuerdo con esta definición, (1), (2) y (3) son condiciones necesarias y suficientes para el conocimiento. Las condiciones (1) y (2) se aceptan generalmente sin demasiada discusión: no es posible conocer una falsedad, y es

necesario creer en lo que se pretende conocer. Además, muy pocos han cuestionado la necesidad de alguna forma de justificación adecuada, de acuerdo con lo que se establece en (3): si creo que Noggin ganará el Derby de Kentucky y apuesto por él en la lista de caballos y jinetes, nadie supondrá que yo lo sabía, ni siquiera si Noggin termina siendo el primero. Simplemente habré tenido suerte.

Gettier lo complicó todo Como cabía prever, la atención se centró principalmente en la forma precisa y en el grado de justificación que requiere la condición (3), pero el esquema básico que proporciona la teoría tripartita se mantuvo vigente durante unos 2 500 años. Justo hasta que, en 1963, el filósofo norteamericano Edmund Gettier lo complicó todo de forma muy hábil. En un breve artículo, Gettier proporcionaba contraejemplos, en un espíritu similar al relato de Don, Eric y Alee, en los cuales alguien elaboraba una creencia simultáneamente verdadera y justificada —es decir, que reunía las tres condiciones estipuladas por la teoría tripartita— y que, no obstante, no permitía considerar como conocimiento lo que ese alguien pensaba que sabía.

El problema que plantean los ejemplos del tipo que ofrece Gettier es que en esos casos la justificación para sostener una creencia no se encuentra relacionada del modo adecuado con la verdad de esa creencia, de modo que su verdad es más o menos una cuestión de suerte. Desde entonces se han hecho muchos esfuerzos para intentar tapar el agujero que mostró Gettier. Algunos filósofos han cuestionado la totalidad del proyecto que pretende definir el conocimiento a partir de las condiciones necesarias y suficientes. Pero, por lo general, las tentativas de resolver el problema de Gettier han consistido en proponer una escurridiza «cuarta condición» que permita reforzar el modelo platónico.

1963

La teoría tripartita del conocimiento

Muchas de las mejoras del concepto de justificación son de naturaleza «externalista», y se centran en factores que se encuentran fuera de los estados psicológicos del conocedor putativo. Por ejemplo, una teoría causal insiste en que el paso de la creencia verdadera al conocimiento depende de que la creencia sea causada por factores externos pertinentes. La creencia de Don no puede considerarse conocimiento porque se encuentra causalmente relacionada con la persona equivocada (Alec, no Eric).

Desde la publicación del artículo de Gettier, la busca de un «parche» se ha desarrollado como una especie de carrera armamentística filosófica. Las tentativas de mejora de la definición tripartita han topado con una batería de contraejemplos destinados a mostrar que persiste algún fallo. Y las propuestas que, en apariencia, esquivan el problema de Gettier tienden a conseguirlo al

precio de marginar una parte importante de lo que intuitivamente consideramos conocimiento.

¿Debería ser indefectible el conocimiento?

Una propuesta para la cuarta condición suplementaria a la teoría tripartita es que el conocimiento debería ser lo que los filósofos llaman «indefectible». La idea es que no debería existir nada de lo que alguien pudiera conocer que invalidara las razones que tiene para creer en ello. Así, por ejemplo, si Don hubiera sabido que Eric tenía un hermano gemelo idéntico, su creencia de que el hombre apoyado contra la farola era Eric no hubiera estado justificada. Pero, de acuerdo con el mismo razonamiento, si el conocimiento tiene que ser indefectible de este modo, Don no hubiera *sabido* que era Eric ni siquiera *si lo hubiera sido*. Esto es así tanto si Don conoce como si no la existencia del hermano gemelo; siempre podría haber un factor de ese tipo, de modo que siempre existirá un sentido en el que quienes saben no saben que saben. Como muchas otras respuestas al problema de Gettier, la demanda de indefectibilidad corre el riesgo de poner el listón tan alto que muy pocas de la cosas que normalmente consideramos conocimiento pasen la prueba.

La comedia de los errores

El recurso a las identidades falsas, en especial a los gemelos idénticos, para cuestionar el conocimiento que, en apariencia, poseemos justificadamente, es una técnica del todo conocida para cualquier persona familiarizada con las obras de Shakespeare. Por ejemplo, en *La comedia de los errores* no hay sólo un par de gemelos idénticos, sino dos: Antífolo y Dromio de Siracusa, y Antífolo y Dromio de Éfeso (a los que un naufragio separó al nacer). Shakespeare los reúne a todos para crear una farsa ingeniosa que puede analizarse como los contraejemplos de Gettier. Así, cuando Antífolo de Siracusa llega a Éfeso, Angelo, el orfebre local, se dirige a él como «mi señor Antífolo». Antífolo, que no había puesto un pie en Éfeso nunca, responde confuso: «Vaya, ése es mi nombre». Angelo responde: «Lo sé perfectamente, señor». Pero de hecho. Angelo no *sabe* nada. De acuerdo con la teoría tripartita, su creencia está justificada, aunque es una pura coincidencia que su cliente tenga un gemelo idéntico con el mismo nombre.

La idea en síntesis: ¿cuándo conocemos realmente?

Desde el siglo XVII el progreso de la ciencia lo ha barrido todo a su paso. El trayecto recorrido por Copérnico, Newton, Darwin y Einstein ha marcado numerosos hitos significativos a lo largo del camino, y alimentado la esperanza de que, con el tiempo, incluso las regiones más remotas del universo, y los secretos más profundos del átomo, serán desvelados por la ciencia. ¿Acaso lo conseguirán? Hay algo —que es al mismo tiempo lo más obvio y lo más misterioso de todo— que se ha resistido siempre, tanto a los esfuerzos de los mejores científicos como a los de los mejores filósofos: la mente humana.

Todos nosotros somos inmediatamente conscientes de nuestra conciencia, conscientes de tener pensamientos, sentimientos, deseos, que son subjetivos y privados; de ser actores situados en el centro de nuestro mundo, y de tener una perspectiva singular y personal del mismo. En claro contraste con esta conciencia, la ciencia es triunfalmente objetiva, susceptible de escrutinio, y se abstiene de lo personal y de la perspectiva. Entonces, ¿cómo es posible concebir la existencia de algo tan extraño como la conciencia en el mundo físico que describe la ciencia? ¿Cómo es posible explicar los fenómenos mentales a partir de, o en relación con, estados y acontecimientos corporales? Estas dos preguntas constituyen el problema del cuerpo y la mente.

Igual que en la epistemología (la filosofía del conocimiento), el francés René Descartes dejó una impronta en la filosofía de la mente del siglo XVII que persistió a lo largo de la historia de la filosofía hasta la actualidad. La opción cartesiana de refugiarse en el propio yo (véase la página 20) lleva de un modo natural a atribuir a la mente un eminente estatus en relación con el mundo exterior. Desde el punto de vista metafísico, Descartes concebía la mente como una entidad completamente distinta: como una sustancia mental cuya naturaleza esencial es el pensamiento. Todo lo demás es materia, o sustancia material, y su característica definitoria es la extensión espacial (es decir, el ocupar un espacio físico). Así, el cartesianismo suponía dos reinos distintos, uno de mentes inmateriales, con propiedades mentales tales como el pensamiento y los sentimientos; y el otro de cuerpos materiales, con propiedades físicas tales como la materia o la forma. Esta representación de la relación entre mente y materia, conocida como « dualismo de las sustancias» es la que Gilbert Ryle denunció como « el dogma del fantasma en la máquina» (véase el cuadro).

Cronología

1637	1644	1655
El problema del cuerpo y la mente	*Cogito ergo sum*	El barco de Perseo

El fantasma de Ryle

En su libro *El concepto de lo mental* (1949) el filósofo inglés Gilbert Ryle argumenta que la cartesiana concepción dualista de la mente y la materia se basa en un « error de categorías» . Imaginemos, por ejemplo, a un turista al que se le enseñaran por separado las distintas facultades, librerías y otros edificios que integran la Universidad de Oxford y que, no obstante, al terminar el paseo, se quejara de no haber visto la universidad. El turista habría atribuido erróneamente a la misma categoría de existencia tanto la universidad como los edificios de los que se compone, deformando así la relación que existe entre ambos. Desde el punto de vista de Ryle, Descartes comete el mismo error garrafal en el caso de la mente y la materia, al suponer de forma equivocada que se trata de dos sustancias completamente distintas. Semejante metafísica dualista es la que inspira la despectiva imagen del « fantasma en la máquina» de Ryle: el alma o la mente inmaterial (el fantasma) que vive en el cuerpo y controla los mandos del cuerpo material (la máquina). Tras presentar su incisivo ataque al dualismo cartesiano, Ryle ofrece su propia respuesta al problema del cuerpo y la mente: el behaviorismo (véase capítulo 10)

Problemas del dualismo El deseo de beber lleva a mi mano a levantar el vaso; una chincheta en el pie me produce dolor. El cuerpo y la mente interactúan (eso indica el sentido común): los acontecimientos mentales dan lugar a los físicos, y viceversa. Pero la necesidad de tal interacción pone inmediatamente en duda la representación cartesiana. Que un efecto físico requiera una causa física es un principio científico fundamental, pero al concebir la mente y la materia como cosas *esencialmente* distintas, Descartes parece haber hecho imposible tal interacción.

1690	1912	1950	1974
El velo de la percepción	Otras mentes	El test de Turing	¿Cómo es ser un murciélago?

«**El dogma del fantasma en la máquina … sostiene que existen cuerpos y mentes, que tienen lugar procesos físicos y procesos mentales, que existen causas**

El propio Descartes advirtió el problema, y que era precisa la intervención de Dios para establecer la necesaria relación causal. El más joven de sus discípulos contemporáneos, Nicolás Malebranche, asumió su dualismo y se propuso resolver él mismo el problema de la causalidad. Su asombrosa

mecánicas para los movimientos corporales y causas mentales para los movimientos corporales.»

Gilbert Ryle, 1949

solución fue afirmar que la interacción no tenía lugar en absoluto. Cada vez en que se requería la conjunción de acontecimientos mentales y físicos, Dios intervenía para que tuviera lugar, y ello producía la apariencia de causa y efecto. Esta débil doctrina, conocida como «ocasionalismo», ha cosechado pocos adeptos, y sirve fundamentalmente para ilustrar la seriedad del problema que se pretendía resolver.

Una manera muy seductora de evitar algunos problemas relativos a la posición cartesiana, es el *dualismo de las propiedades*, propuesto en la obra de un contemporáneo holandés de Descartes, Baruch Spinoza, quien sostenía que la noción de dualismo se refiere no a las sustancias sino a las propiedades: dos tipos distintos de propiedades, mentales y físicas, pueden atribuirse a una misma cosa (una persona o un objeto), pero estos atributos son irreductiblemente diferentes y los términos para analizarlos no son intercambiables. Así pues, las diferentes propiedades describen distintos *aspectos* de la misma entidad (de ahí que, en ocasiones, se llame a esta perspectiva «teoría del doble aspecto»). La teoría permite explicar cómo tiene lugar la interacción mente-cuerpo, puesto que las causas de nuestras propias acciones tienen tanto aspectos mentales como físicos. Pero al atribuir propiedades esencialmente tan distintas a un mismo objeto, surge la sospecha de que el dualismo de las propiedades se haya limitado a desplazar el principal problema que debe afrontar el dualismo de la sustancia, en vez de resolverlo.

El fisicalismo La respuesta evidente a los problemas con que topa el dualismo de la sustancia cartesiano es la adopción de una posición monista, la afirmación de que existe una única «cosa» en el mundo, sea física o mental, y no dos. Algunos autores, el más notable de ellos George Berkeley (véase la página 18), se han decantado por el idealismo, sosteniendo que la realidad consiste únicamente en las mentes y sus ideas. Pero la gran mayoría de los filósofos, más aún entre los actuales, han optado por alguna forma de explicación fisicalista. Influido por los indiscutibles éxitos de la ciencia en otros campos, el fisicalista insiste en que también la mente debe inscribirse en el ámbito de la ciencia; y puesto que los objetos de estudio de la ciencia son exclusivamente físicos, también la mente debe ser física. La tarea consiste, entonces, en aclarar cómo encaja la mente (subjetiva y privada) en una explicación completa y puramente física del mundo (objetivo y universalmente accesible).

El fisicalismo ha adoptado muchas formas distintas, pero todas ellas tienen en común el ser reductivas: pretenden mostrar que el fenómeno de lo mental

Los orígenes del dualismo

Descartes fue el primero en

puede analizarse de manera completa y exhaustiva, en términos puramente físicos. Los avances en neurociencia evidencian con bastante claridad que los estados mentales se encuentran íntimamente vinculados con los estados cerebrales. Así, la vía simple para el fisicalista es afirmar que los fenómenos mentales son idénticos a los hechos y a los procesos cerebrales. Las versiones más radicales de tales teorías de la identidad son «excluyentes»: proponen que, del mismo modo que ocurre con los avances de nuestra comprensión científica, la «psicología popular» —la manera habitual de expresar y pensar nuestra vida mental, en términos de creencias, deseos, intenciones…— desaparecerá y será sustituida por conceptos precisos y descripciones procedentes principalmente de la neurociencia.

brindar el clásico planteamiento del dualismo de la sustancia, pero no fue en absoluto el primer dualista. De hecho, hay un cierto dualismo implícito en cualquier filosofía, religión o cosmovisión que presuponga que existe un reino sobrenatural en el que residen cuerpos inmateriales (almas, dioses, demonios, ángeles y cosas similares). La idea de que un alma puede sobrevivir a la muerte del cuerpo físico o reencarnarse en otro cuerpo (humano o de otra especie) también implica un determinado tipo de concepción dualista del mundo.

Las soluciones fisicalistas al problema del cuerpo y la mente hacen caso omiso a muchos de los aspectos problemáticos del dualismo. En particular, los misterios causales que atormentan a los dualistas se desvanecen por el simple hecho de inscribir la conciencia en el ámbito de la explicación científica. Naturalmente, los críticos del fisicalismo lamentan que sus representantes dejen de lado tantas cosas; que sus éxitos se consigan a un precio tan caro (a fuerza de renunciar a captar la esencia de la conciencia en cuanto experiencia, su naturaleza subjetiva).

La idea en síntesis: demasiado para mi mente

08 ¿Cómo es ser un murciélago?

«Imagina que tuviéramos unas membranas en los brazos que nos permitieran volar al caer el sol y cazar insectos con la lengua al alba; que tuviéramos un vista muy limitada y percibiéramos el mundo circundante mediante un sistema de emisión de signos de alta frecuencia sonora; y que pasáramos el día colgados cabeza abajo en el techo de un ático. Hasta donde alcanza mi imaginación (que no es mucho), esto sólo me dice cómo sería para mí comportarme como un murciélago. Pero no se trata de eso. Lo que quiero saber es cómo es ser un *murciélago* para un murciélago.»

En la filosofía de la mente el artículo del filósofo norteamericano Thomas Nagel escrito en 1974 y titulado «¿Cómo es ser un murciélago?» ha sido quizás más influyente que cualquier otro artículo publicado recientemente. Nagel recogió de forma sucinta la esencia del malestar que a muchos filósofos les infundían las tentativas recientes de análisis de la vida mental y de la conciencia en términos puramente físicos. Ello explica que su artículo se haya convertido en casi un tótem para los filósofos insatisfechos con semejantes explicaciones fisicalistas, y con las teorías de la mente reduccionistas.

La perspectiva del murciélago El principal argumento de Nagel es que existe el «carácter subjetivo de la experiencia» —una cosa es *ser* un organismo particular, otra es serlo, de un modo determinado, *para* ese organismo— de lo que nunca se da cuenta en tales explicaciones reduccionistas. Consideremos el caso de un murciélago. Los murciélagos vuelan y cazan insectos en la oscuridad gracias a un sistema de ultrasonidos, de ecolocalización, según el cual emiten chillidos de alta frecuencia y detectan sus ecos tras rebotar en la superficie de los objetos circundantes. Esta forma de percepción es completamente distinta a cualquiera de nuestros sentidos, de modo que es razonable suponer que se trata de una subjetividad completamente distinta de cualquier cosa que seamos capaces de experimentar. En efecto, existen experiencias que como humanos nunca podremos experimentar; existen hechos de la experiencia cuya naturaleza exacta está fuera del alcance de nuestra comprensión. La incomprensión esencial de estos hechos se debe a su naturaleza subjetiva, al hecho de que, en esencia, entrañan un punto de vista singular.

Cronología

c. 250 a. C.	1637	1655
¿Los animales sienten dolor?	El problema del cuerpo y la mente	El barco de Teseo

Entre los filósofos fisicalistas existe una tendencia a

«Sin la conciencia el

citar ejemplos de reducciones científicas eficaces —por ejemplo, el análisis del agua como H_20, o de la luz como descarga eléctrica—, y a sugerir a partir de estos ejemplos que existen casos similares para el análisis de los fenómenos mentales considerados como fenómenos físicos. Nagel discute esta posibilidad: el éxito de este tipo de análisis científico consiste en conseguir una gran objetividad a fuerza de abandonar el punto de vista subjetivo: y precisamente la omisión de este elemento subjetivo en las teorías fisicalistas de la mente las convierte en incompletas e insatisfactorias. Nagel concluye que « el modo en que podría revelarse el verdadero carácter de las experiencias en las operaciones físicas es un misterio» , y eso es todo lo que la ciencia nos ofrece.

problema del cuerpo y la mente sería mucho menos interesante. Con la conciencia parece no permitirnos albergar ninguna esperanza.»

Thomas Nagel, 1979

Lo que Mary no sabe Nagel parece contentarse con dejar el asunto como un misterio, con evidenciar la incapacidad de las teorías fisicalistas recientes para dar cuenta del elemento subjetivo que parece esencial a la conciencia. Además, afirma oponerse sólo a los planteamientos reduccionistas, no al fisicalismo como tal. Pero el filósofo australiano Frank Jackson pretende ir más lejos. En un artículo muy polémico del año 1982 titulado « Lo que Mary no sabe» , plantea un experimento mental sobre una chica que conoce cualquier estado físico imaginable sobre el color. Pues bien, si el fisicalismo estuviera en lo cierto, arguye Jackson, Mary conocería todo lo que hay que conocer. Pero resulta que hay cosas (hechos) que no conoce en absoluto: no sabe cómo es ver los colores; ha aprendido cómo es ver el rojo (etcétera). Jackson concluye que existen hechos que no son, y no pueden ser, captados por una teoría física (los que no son físicos), y en consecuencia que el fisicalismo es falso (véase el cuadro).

1912	1953	1974
La mente de los otros	El escarabajo en una caja	¿Cómo es ser un murciélago?

Mary en blanco y negro

Desde que nació, Mary se encontraba confinada en una habitación en blanco y negro, en la que nunca se la había expuesto a nada que no fuera negro, blanco o de algún otro tono en la escala de grises. Recibió efectivamente educación, a pesar de ser muy peculiar, y mediante la lectura de libros (sin colores, por supuesto) y el visionado de conferencias en blanco y negro en una televisión, terminó convirtiéndose en una de las mayores científicas mundiales. Aprendió literalmente todo lo que había que aprender (y todo lo que hasta entonces se hubiera

podido conocer) sobre la naturaleza física del mundo, sobre nosotros y nuestro entorno. Y, por fin, llegó el día en que se liberó a Mary de su habitación monocroma dejándola salir al mundo exterior. ¡Qué impresión le causó! Veía los colores por primera vez. De pronto supo cómo era ver el rojo, el azul, el amarillo. De modo que, a pesar de conocer todos los hechos físicos relativos al color, había cosas sobre el color que no sabía.

Moraleja:

1) hay hechos que no son físicos;

2) ten cuidado con los padres que escoges.

Naturalmente el argumento de Jackson no ha persuadido a los físicalistas convencidos. Las objeciones se dirigen en primer lugar al estatus de lo que él llama « hechos que no son físicos» : algunos críticos aceptan que son hechos pero niegan que no sean físicos, mientras que otros afirman rotundamente que no son hechos. El origen de estas objeciones generalmente se debe a que Jackson se opone al fisicalismo de antemano: si el fisicalismo es cierto, y Mary conoce todos los hechos físicos que quepa conocer acerca del color, entonces conocerá por añadidura todo lo que cabe conocer sobre la rojez, incluidas las experiencias subjetivas asociadas a esa condición. Además, el modo en que Jackson utiliza los estados psicológicos de Mary para establecer la distinción necesaria entre los hechos físicos y los que no lo son hace sospechar que el argumento incurre en la « falacia del hombre de la máscara» (véase el cuadro de la página siguiente).

El hombre de la máscara

De acuerdo con la ley leibniziana de la « identidad de los indiscernibles» , si dos cosas, A y B, son idénticas, entonces cada propiedad de A será también una propiedad de B; y por lo mismo, si A tiene una propiedad de la que carece B, entonces A y B no son cosas idénticas. Pues bien, Bertie piensa que Bono es la estrella del rock más brillante que existe en el mundo, pero no tiene opinión acerca de Paul Hewson (no sabe que, de hecho, ése es el nombre original de Bono). Ello implica que Bono posee una propiedad de la que carece Paul Hewson —la propiedad de ser considerado como la estrella del rock más brillante del mundo— y, por lo tanto, de acuerdo con la ley de Leibniz, Bono no es la misma persona que Paul Hewson. Pero sí lo es, de modo que algo falla en el argumento. El

problema se debe a la llamada « falacia del esencialmente hombre de la máscara» : ignoro quién es el hombre de la máscara; sé quién es mi hermano; por lo tanto mi hermano no es el hombre de la máscara. El origen de la falacia se encuentra en el hecho de que los puntos de vista subjetivos, o nuestras creencias, no constituyen propiedades reales de una cosa: podemos conocer algo a partir de distintas descripciones, en la medida en que distintas personas pueden creer cosas distintas sobre la misma cosa.

Ahora bien ¿el experimento mental de Frank Jacksonincurre en la falacia del hombre de la máscara? Los críticos fisicalistas sostienen que Jackson usa un « tipo equivocado» de propiedad para establecer una dualidad entre los hechos físicos y los que no lo son. Pero en realidad, prosiguen, sólo existe un tipo de hechos (los físicos) sobre los que pueden darse distintas descripciones, subjetivas: dos formas distintas de mirar lo mismo.

Con independencia de la fuerza de los argumentos contra Mary, resulta difícil evitar la impresión de que tanto Jackson como Nagel metieron el dedo en *algún sitio* (o de que, en las versiones que el fisicalismo ha propuesto hasta la fecha, falta algo esencial). Tal vez sólo pueda concluirse que el propósito de remitir la conciencia a una explicación meramente física del mundo todavía tiene mucha vida por delante.

La idea en síntesis: ¿la mente de un murciélago?

Parry: Los polis detienen a la gente equivocada. Entrevistador: A veces detienen a los que toca.
Parry: La ley debería pillar a los chorizos de los italianos. Entrevistador: ¿Usted es italiano?
Parry: Soy de origen británico, pero nací en este país. Entrevistador: ¿Conoce usted a chorizos italianos?
Parry: Intento evitar los bajos fondos.
Entrevistador: ¿Le dan miedo los bajos fondos?
Parry: Los de la maña son los peores gángsters. Entrevistador: ¿Tiene relaciones con la mafia?
Parry: Digamos que me conocen.
Entrevistador: ¿Cómo le conocen?
Parry: Me espían.
Entrevistador: ¿Por qué?
Parry: Quieren conseguir pruebas falsas para incriminarme. Entrevistador: ¿Qué tienen contra usted?
Parry: Me están esperando fuera.

Se diría que Parry está en apuros ¿no es cierto? Tal vez, pero su principal problema no es la mafia, sino… el hecho de ser un programa informático. Creado en 1971 por un psiquiatra de la Universidad de Stanford, Kenneth Colby, Parry fue creado para responder a preguntas del modo en que lo hace un esquizofrénico que padece la fijación paranoica de estar en el punto de mira de la mafia.

A Parry lo interrogaron junto con unos cuantos pacientes auténticamente paranoicos y los resultados fueron luego evaluados por un equipo de psiquiatras. Ningún médico de ese equipo adivinó que Parry no era un paciente real.

Cronología

c. 1637	1912
El problema del cuerpo y la mente	Otras mentes

¿Es posible que Parry piense? Veinte años antes del nacimiento de Parry, en 1950, el matemático inglés pionero en la informática, Alan Turing, escribió un artículo capital en el que proponía un test para determinar si era posible que una máquina pensara. El test, basado en un juego de grupo que se llamaba el juego de imitación, requería un interrogador que se comunicase con otro humano y con una máquina, ambos físicamente separados del interrogador, mediante una especie de conexiones electrónicas. El interrogador no podía hacer ninguna

pregunta para diferenciar a la máquina del humano, y si tras un determinado período de tiempo era incapaz de distinguirlos se suponía que la máquina había superado el test.

¿Pasó Parry el test? En realidad, no. Para funcionar como un test de Turing adecuado, el equipo de psiquiatras (que aquí hacen el papel del interrogador) debería haber sabido que uno de los pacientes era de hecho un ordenador y que la tarea era identificarlo. En cualquier caso, Parry se hubiera delatado enseguida si hubiera sido sometido a un interrogatorio más largo. El propio Turing creía que a finales del siglo XX los avances en la programación informática habrían alcanzado un punto en el que no existiría más que un 70% de posibilidades de identificar correctamente la diferencia después de cinco minutos de entrevista, aunque de hecho el progreso haya ido bastante más despacio de lo que él predijo. Hasta el momento ningún programa informático ha estado cerca de superar el test de Turing.

«**Creo que a finales de siglo [XX] el uso de las palabras y la opinión general de la gente educada se habrá transformado tanto que seremos capaces de hablar de máquinas pensantes sin que se nos contradiga.**»

Alan Turing, 1912-1954

Turing propuso su test para evitar la pregunta « ¿Las máquinas pueden pensar?» , que consideraba demasiado imprecisa como para que mereciera la pena contestar, pero el test se acepta hoy comúnmente como un criterio mediante el que juzgar si un programa es capaz de pensar (o « tiene mente» o « muestra inteligencia» de acuerdo con la prueba). Como tal, los defensores (científicos y filosóficos) de una « IA (inteligencia artificial) fuerte» —la tesis de que las computadoras debidamente programadas poseen mente (no sólo son simulaciones de una mente) en el mismo sentido en que la tienen los humanos— consideran el test como un punto de referencia.

1950	1974
El test de Turing	¿Cómo es ser un murciélago?

La habitación china El filósofo norteamericano John Searle planteó en 1980 un experimento mental que supuso el desafío más decisivo al test de Turing. En él, Searle se imaginaba a sí mismo —un hablante inglés que no sabe una palabra de chino— confinado en una habitación en la que le pasaban textos escritos por tandas. Le habían proporcionado un montón de símbolos chinos y una voluminosa guía en inglés en la que se explicaba cómo tenía que combinar determinados símbolos para devolver la respuesta a las series de símbolos que le mandaban por tandas. Con el tiempo adquirió tal destreza en la tarea que, desde el punto de vista de alguien del exterior, sus respuestas eran indistinguibles de las de un hablante nativo de chino. Es decir, el toma y daca entre el interior y el exterior de la

habitación era exactamente igual que si hubiera existido una plena comprensión del chino. No obstante, lo único que él estaría haciendo sería manipular signos formales ininteligibles, pues no entendía nada.

> **«Al final se reconocerá que las actuales tentativas de entender la mente por analogía con las computadoras artificiales que pueden interpretar magníficamente algunas de las mismas tareas externas que los seres conscientes son una pérdida de tiempo garrafal.»**
>
> **Thomas Nagel, 1986**

Producir mensajes correctos en respuesta a los que recibe, de acuerdo con las reglas que le proporciona un programa (equivalente a la guía en inglés de Searle), es exactamente lo que hace una computadora. Searle sugiere que, del mismo modo que el individuo de la habitación china, un programa informático, aunque más sofisticado, no es más —y nunca puede ser más— que un manipulador de signos desprovisto de mente; es esencialmente sintáctico —se atiene a unas reglas para combinar símbolos— pero no puede comprender el significado, la semántica. Del mismo modo que en la habitación china no se produce ninguna comprensión, tampoco se produce en el programa informático: no hay comprensión, ni inteligencia, ni mente; lo único que se da es una simulación de todos estos procesos.

Superar el test de Turing es fundamentalmente cuestión de proporcionar las respuestas adecuadas a los mensajes recibidos, de modo que, a la luz de la habitación china, la pretensión del test de permitir determinar si una máquina piensa queda minada. Y con la caída del test de Turing, también se hunde la tesis central de la IA fuerte. Pero éstas no son las únicas bajas. Existen dos aproximaciones más a la filosofía de la mente que quedan invalidadas si se acepta el planteamiento de la habitación china.

En la cultura popular

Arthur C. Clarke se creyó a pies juntillas la predicción de Alan Turing. Cuando en 1968 colaboró con Stanley Kubrick en *2001. Una odisea en el espacio* creó una computadora inteligente que se llamaba HAL (cada una de las letras es la anterior a la de las siglas de IBM). En la película, a ninguno de los humanos le sorprende que una máquina pensante dirija su viaje en el espacio.

Problemas del behaviorismo y del funcionalismo La idea central del behaviorismo consiste en que los fenómenos mentales pueden traducirse, sin que

se pierda nada de su contenido, en tipos de comportamiento o en disposiciones a determinados comportamientos. Así, decir que alguien siente dolor es una forma abreviada de decir que está sangrando, retorciéndose, etc. Es decir, los hechos mentales se definen en términos completamente externos, como registros (*inputs*) y respuestas (*outputs*) observables, el valor de las cuales es negado explícitamente por la habitación china. Desde la clásica exposición que brindó Gilbert Ryle (véase página 33), el behaviorismo ya había sido sometido a innumerables objeciones antes de la aparición de Searle. Su importancia actual se debe sobre todo a que fecundó una doctrina que es la teoría de la mente más unánimemente aceptada: el funcionalismo.

El funcionalismo, que resuelve algunos de los problemas del behaviorismo, afirma que los estados mentales son estados funcionales: un determinado estado mental se identifica como tal en virtud del papel o de la función que tiene en relación con varios registros (*inputs*) (sus típicas causas), los efectos que produce en otros estados mentales y las varias respuestas (*outputs*) (sus típicos efectos en los comportamientos). Si usamos una analogía informática, el funcionalismo (como el behaviorismo) es una « solución de software» a la teoría de la mente: define los fenómenos mentales como registros (*inputs*) y respuestas (*outputs*), sin tener en cuenta el hardware (dualista, fisicalista, o cualquier otro) en el que opera el software. Naturalmente, el problema es que analizar los registros y las respuestas parece abocamos a recaer en la habitación china.

La idea en síntesis: ¿te has sometido al test alguna vez?

10 El barco de Teseo

«¡Caramba, cuántos problemas está dando a Teo el coche que compró a Joe! Empezó con algunos problemas: había que cambiar la cerradura de una puerta, alguna dificultad en la suspensión trasera, lo normal. Luego empezaron a fallar cosas más gordas: primero el embrague, después la caja de cambios, y al final todo el sistema eléctrico. Y por el camino hubo muchas otras averías, de modo que el coche raramente salía del taller. Y así iba, una y otra vez… Increíble. "Lo más increíble —pensó Teo apesadumbrado— es que el coche sólo tiene dos años y le he cambiado *cada una de las piezas*. Anda, mira el lado bueno del asunto… ¡va a resultar que tengo un coche nuevo!"»

¿Acaso tiene razón Teo? ¿O sigue siendo el mismo coche? La historia de Teo —o del barco de Teseo— es uno de los muchos puzles que han utilizado los filósofos a lo largo de los tiempos para someter a prueba nuestras intuiciones acerca de la identidad de cosas o personas. Parece que nuestras ideas en este terreno suelen ser tan firmes como problemáticas. Thomas Hobbes fue quien relató y desarrolló la historia del barco de Teseo. Si volvemos a la versión de Teo… « Joe *el honesto* no había estado a la altura de su nombre. La mayoría de las piezas que había cambiado en el coche de Teo funcionaban bien, e incluso arregló algunas que estaban mal. Había conservado las partes viejas y las había encajado unas con otras. Al cabo de dos años, había armado una copia exacta del coche de Teo. Pensaba que era una copia. ¿Acaso era el coche de Theo?»

Las crisis de identidad ¿Cuál es el original? ¿El coche de Teo, hecho hoy a base de nuevas piezas, o la versión de Joe, compuesta con las piezas originales? Posiblemente depende de a quién se le pregunte. Pero la identidad del coche ha sido a lo largo de la historia mucho menos clara y distinta de lo deseable. El problema no se limita a los coches o a los barcos. La gente cambia muchísimo a lo largo de una vida. Tanto física como psicológicamente, debe haber muy pocas cosas en común entre un niño de 2 años y el abuelo de 90 que ha ocupado su lugar 88 años más tarde. ¿Son pues la misma persona? Si es así, ¿qué hace que lo siga siendo? El asunto no es irrelevante: ¿es justo castigar a un hombre de 90 años por algo que hizo 70 años atrás? ¿Y si no lo recuerda? ¿Debería un doctor dejar morir al anciano de 90 años si una versión (supuestamente) anterior suya hubiera expresado ese deseo 40 años atrás?

Cronología

1637	1644
El problema del cuerpo y la mente	*Cogito ergo sum*

Éste es el problema de la identidad personal, que ha mantenido en vilo a los

filósofos durante cientos de años. ¿Cuáles son las condiciones necesarias y suficientes para que una persona sea en un determinado momento la misma persona que un tiempo después?

Los animales y los trasplantes de cerebro Desde el punto de vista del sentido común posiblemente la identidad personal sea un problema biológico: soy ahora el mismo que en el pasado porque soy el mismo organismo vivo, el mismo animal humano; estoy unido a un cuerpo determinado que es una entidad orgánica única y continua. Pero imaginemos por un momento un trasplante de cerebro —una operación que cabe imaginar al alcance de la tecnología futura— en la que colocan *tu* cerebro en mi cuerpo. Nuestra intuición seguramente indica que *tú* tienes un cuerpo nuevo, no que mi cuerpo tenga un cerebro nuevo; si es así, parece que tener un cuerpo determinado *no* es una condición necesaria de la subsistencia de la persona.

Esta consideración ha llevado a muchos filósofos a renunciar al cuerpo a favor de la mente: la afirmación de la identidad está vinculada no a la totalidad del cuerpo sino tan sólo al cerebro. Este desplazamiento se adecúa a nuestra intuición acerca del trasplante de cerebro pero no resuelve del todo el problema. Nuestra preocupación es qué suponemos que emana del cerebro, no el órgano físico en sí mismo. Mientras que podemos seguir dudando de cómo surge la conciencia o la actividad mental de la actividad del cerebro, pocos dudamos en cambio de que el cerebro es el origen de esa actividad. Al considerar qué hace que yo sea yo, lo que me preocupa es el « software» de experiencias, recuerdos, creencias, etc., no el « hardware» de una determinada masa de materia gris. Mi sentido de ser yo no se vería demasiado alterado si copiaran la totalidad de mis experiencias, recuerdos, etc., en un cerebro sintético, o si el cerebro de alguna otra persona pudiera reconfigurarse para acoger todos mis recuerdos, mis creencias, etc. Yo soy mi mente: voy donde mi mente va. Según esto, mi identidad no se encuentra en absoluto asociada a mi cuerpo físico, ni siquiera al cerebro.

1655	1950	1981
El barco de Teseo	El test de Turing	El cerebro en una cubeta

El oficial valiente

Thomas Reid intentó socavar la posición de Locke con la siguiente historia: « En su época de escolar, un valiente oficial había recibido unos azotes como castigo por robar en un huerto; en su primera campaña militar consiguió apoderarse del estandarte del enemigo; al final de su vida fue nombrado general. Supongamos que cuando se apoderó del estandarte todavía recordaba los azotes, pero que cuando lo nombraron

general recordaba cómo se había apoderado del estandarte pero ya no de haber sido azotado».

Locke podría aceptar las consecuencias de la objeción de Reid: su tesis implica una clara distinción entre el ser humano (el organismo) y la persona (el sujeto de la conciencia), de modo que el anciano general sería en un sentido real una persona distinta del joven.

La continuidad psicológica Adoptemos una perspectiva psicológica para abordar la cuestión de la identidad personal, en vez de la biológica o la física, y supongamos que cada parte de mi historia psicológica se encuentra unida a otras partes anteriores por hilos de recuerdos, creencias, etc., perdurables. No todos ellos (quizá ninguno) deben extenderse desde el principio hasta el final; mientras exista uno solo que solape todo el entramado de tales elementos, mi historia se mantendrá. Seguiré siendo yo. La idea de la continuidad psicológica como el principal criterio de identidad personal a lo largo del tiempo es de John Locke. Sigue siendo la teoría hegemónica entre los filósofos contemporáneos, aunque no está exenta de problemas.

Imaginemos, por ejemplo, un sistema de teletransportación como el de *Star Trek*. Supongamos que tal sistema registra nuestra composición física hasta el último átomo y transfiere luego estos datos a algún lugar remoto (por ejemplo desde Londres en la Tierra hasta la base número 1 de la Luna), donde nuestro cuerpo es replicado exactamente (con materia nueva) en el preciso instante en el que nuestro cuerpo es aniquilado en Londres. Todo es perfecto, siempre que se adhiera la tesis de la continuidad psicológica: existe un flujo de recuerdos, etc., que fluye desde el individuo en Londres hasta el que está en la Luna, de modo que la continuidad psicológica y, por lo tanto, la identidad personal se preservan. Estamos en la base número 1 de la Luna. Pero supongamos que el sistema falla y olvida realizar la aniquilación que corresponde en Londres. Ahora existen dos «yo»: uno en la Tierra y otro en la Luna. De acuerdo con la explicación de la continuidad, puesto que el flujo se preserva en los dos casos, existen dos yo. En este caso, no dudaremos mucho en decir que somos el individuo que está en Londres mientras que el que está en la Luna es una copia. Pero si esta intuición es cierta, ello parece obligamos a volver a la explicación biológica/animal: se diría que somos el antiguo cuerpo que se encuentra en Londres y no el nuevo que hay en la Luna.

Aclaremos el yo Semejante contradicción entre las distintas intuiciones podría deberse a que hacemos las preguntas equivocadas, o a que aplicamos conceptos erróneos para responderlas. David Hume se ocupó del carácter elusivo del yo, y

sostuvo que, por más que observemos atentamente en nuestro interior, sólo podemos detectar pensamientos, recuerdos, experiencias aisladas. Aunque resulta natural imaginar un yo sustancial como sujeto de tales pensamientos, Hume sostiene que tal idea es errónea: el yo no es más que el punto de vista que permite que nuestros pensamientos y nuestras experiencias tengan sentido, pero no puede aprehenderse por sí mismo.

La idea del yo como «cosa» sustancial, nuestra esencia, provoca confusión cuando imaginamos que podremos realizar trasplantes de cerebro o ser aniquilados y reconstruidos en otra parte. En los mencionados experimentos mentales asumimos nuestra subsistencia personal como si de algún modo dependiera de encontrar el lugar en el que se emplaza el yo. Pero si dejamos de pensar el yo como una sustancia, las cosas resultan más claras. Supongamos, por ejemplo, que las funciones del teletransportador aniquilaran correctamente nuestro cuerpo en Londres pero produjeran dos copias en la Luna. Preguntar cuál de los dos soy yo (o «¿dónde ha ido a parar mi yo?») es preguntar mal. El resultado es que tenemos ahora dos seres humanos, cada uno de los cuales parte exactamente del mismo caudal de pensamientos, experiencias y recuerdos; pero cada uno de ellos hará su propio camino y sus historias psicológicas divergirán. Tú (esencialmente el caudal de pensamientos, experiencias y recuerdos) habrás sobrevivido en cada uno de estos dos nuevos individuos: es una forma interesante de subsistencia personal ¡aunque al precio de la propia identidad personal!

La idea en síntesis: ¿qué hace que tú seas tú?

«Todos esos trucos de Hollywood son una perfecta estupidez. La expresión gélida, la mirada de besugo, fija… perfectas chorradas. En realidad, reconocer a un zombi es realmente complicado. Tienen el mismo aspecto que tú o que yo, andan igual, hablan igual (no hay el menor indicio de que no tengan nada en la cabeza). Si le das una buena patada en la espinilla a un zombi verás que pone cara de dolor y chilla cada vez un poquito más fuerte exactamente igual que tú o yo. Pero a diferencia de lo que nos pasa a ti o a mí, no siente nada: ni dolor, ni sensaciones, ni asomo de conciencia. De hecho, digo "tú y yo" pero debería decir sólo "yo". No estoy muy seguro de que tú… cualquiera de vosotros, no resulte ser un zombi.»

«Los zombis en la filosofía se parecen más a las mujeres de la película *Las mujeres perfectas* que a los devoradores de cerebros picados de *La noche de los muertos vivientes*. Aun así, se advierte, aunque sutilmente, algo que no funciona en las mujeres perfectas.»

Larry Hauser, 2006

Los zombis son invitados habituales en el dilatado debate filosófico conocido como « el problema de las otras mentes» . Yo sé lo que tengo en la mente, una vida interior de experiencia consciente, pero el contenido de tu mente es privado y permanece oculto para mí; lo único que puedo observar directamente es tu comportamiento. ¿Acaso eso constituye una evidencia suficiente en la que basar mi idea de que tienes una mente como la mía? Para aliñarlo un poco: ¿cómo sé yo que no eres un zombi como los que hemos descrito, exactamente igual que yo en lo que se refiere al comportamiento y a la psicología, pero aun así desprovisto de conciencia?

Podría parecer absurdo preguntar si los otros tienen mente, ¿pero tal vez haya algo *irracional* en hacerlo? Dada la considerable dificultad para explicar o tomar en cuenta la conciencia en un mundo físico (véase página 36), ¿acaso no es perfectamente racional suponer que la única mente que conozco —la mía— es una excepción, o incluso que es la única? ¿Es posible que todos vosotros —los zombis— seáis normales y yo sea el único raro?

Cronología

c. 250 a. C.	*c.* 350 a. C.	1300
¿Los animales sienten dolor?	Formas del argumento	La navaja de Ockham

Los zombis no son los únicos invitados a las conferencias sobre filosofía de la mente. También encontraréis mutantes. Igual que los zombis, los mutantes filosóficos son menos terroríficos que sus versiones de Hollywood. Por lo demás, resulta imposible distinguirlos de la gente corriente a partir del comportamiento y del aspecto físico. ¡E incluso tienen aunque mente! El problema con los mutantes es que sus mentes no funcionan exactamente del mismo modo que las nuestras (en fin, que la mía seguro).

Los mutantes pueden ser inmensamente distintos: puede darles placer algo que a mí me produciría dolor, puede parecerles rojo algo que yo veo azul; la única regla es que las sensaciones y otros hechos mentales de los mutantes son *diferentes* de los míos. Los mutantes son particularmente útiles cuando se trata de considerar un determinado aspecto del problema de las otras mentes: la duda no es ya si los otros tienen mente, sino si su mente funciona del mismo modo que la mía. ¿Acaso puedo decir, sea en principio, que sientes el dolor igual que yo? Con estas preguntas se abre una zona nueva de debate; y del mismo modo que ocurre con otros aspectos del problema de la mente de los otros, nuestras respuestas contribuyen a elucidar nuestras concepciones básicas de lo que es la mente.

Puesto que somos tan similares en otros sentidos… El modo más común de abordar el problema de las otras mentes, que desarrolló Bertrand Russell entre otros autores, ha consistido en variantes del llamado argumento por analogía. Sé por mí mismo que pisar una chincheta suele ir seguido de determinados tipos de comportamiento (gritar « ¡Ay!» , hacer una mueca de dolor, etc.) y acompañado de una sensación particular: el dolor. De modo que, cuando otra gente se comporta de un modo similar en respuesta a un estímulo similar, puedo inferir que también ellos sienten dolor. De un modo más general, observo innumerables similitudes, tanto psicológicas como del comportamiento, entre yo mismo y los demás, y de estas similitudes concluyo que los demás son asimismo similares con respecto a su psicología.

1637	1912	1950	1953	1974
El problema del cuerpo y la mente	Otras mentes	El test de Turing	El escarabajo en la caja	¿Cómo es ser un murciélago?

El sentido común del argumento de la analogía tiene un aura seductora. En el improbable caso de que se nos diera la palabra para defender nuestra creencia en la existencia de la mente de los otros, probablemente propondríamos alguna

forma de justificación de ese tipo. El argumento es inductivo (véase la página 112), de modo que no puede (y no lo pretende) proporcionar una prueba conclusiva, pero también es cierto que muchos de nosotros nos sentimos justificados para creer en él.

Arremeter contra molinos de viento

Cuando confrontamos el problema de las otras mentes nos parece un caso típico (no el único, según el parecer popular) de filósofos buscando un problema donde a los demás ni siquiera se nos había ocurrido mirar. Es cierto que todos nosotros (incluso los filósofos, aunque sea por razones prácticas) damos por descontado que los otros disfrutan de una vida interior de pensamientos y sentimientos muy parecida a la nuestra. Pero desechar el problema filosófico basándonos en esto es cometer un error. Nadie pretende persuadir a nadie de que los otros sean efectivamente zombis. Se trata más bien de que las distintas maneras en que pensamos en la mente de los otros y en su relación con los cuerpos dan lugar a la *posibilidad* de los zombis. Y ello debería inducirnos a examinar seriamente nuestras concepciones de la mente.

El dualismo cartesiano (véase la página 32) abre una ancha brecha entre los hechos mentales y los físicos, y del abismo resultante surge el escepticismo acerca de las otras mentes. Ésta es una buena razón para examinar críticamente el dualismo, bien en Descartes o bien en sus diversas manifestaciones religiosas. En cambio, uno de los atractivos de las explicaciones fisicalistas de la mente es que permite explicar de forma satisfactoria, en principio por lo menos, los hechos mentales como hechos físicos; y si lo mental se disuelve en lo físico, los zombis se desvanecen simultáneamente. Ello no convierte estas explicaciones en necesariamente ciertas, pero sí evidencia que están bien encaminadas. En este sentido prestar atención al problema de las otras mentes puede arrojar luz sobre otras cuestiones más generales que atañen a la filosofía de la mente.

«Si la relación entre tener un cuerpo humano y un determinado tipo de vida mental es tan contingente como parece indicar la explicación cartesiana de la mente, debería resultarme tan sencillo … concebir una mesa que sufre como lo es concebir el sufrimiento de otra persona. La cuestión, naturalmente, es que

no ocurre así.»

Ludwig Wittgenstein, 1953

La crítica habitual al argumento es que implica una inferencia o una extrapolación de una instancia singular (mi propia mente). Imaginemos, por ejemplo, que encontramos una ostra con una perla e inferimos de ello que todas las ostras tienen perla. Para reducir el riesgo de este tipo de errores, es necesario inspeccionar unas cuantas ostras, pero precisamente ésa es la medida que no podemos adoptar en el caso de las otras mentes. Tal y como señala Wittgenstein: « ¿Cómo puedo generalizar el caso único de un modo tan irresponsable?» .

La irresponsabilidad de sacar conclusiones sobre la base de una sola instancia queda mitigada si la inferencia se hace en el contexto de un cúmulo de información relevante. Por ejemplo, si reconocemos que una perla carece de una utilidad en el funcionamiento de la ostra, o que el valor comercial de las perlas no se corresponde con el hecho de que se las encuentre en todas las ostras, tenderemos menos a inferir erróneamente a partir de un espécimen singular.

El problema con la mente y la conciencia es que resultan tan misteriosas que, a diferencia de lo que ocurre con cualquier otra cosa que nos resulte familiar, no está nada claro qué podría valer como información relevante acumulada. En esta medida el problema de las otras mentes puede considerarse como otro síntoma del problema más general del cuerpo y la mente. Si nuestra teoría de la mente consigue desmitificar la relación entre los fenómenos físicos y mentales, cabe esperar que nuestra preocupación por las otras mentes disminuya o se desvanezca por completo (véase el cuadro de la página anterior).

La idea en síntesis: ¿hay alguien ahí?

12 La guillotina de Hume

«En todos los sistemas morales con los que hasta ahora me he topado, he observado que el autor procede durante un rato de acuerdo con el modo habitual de razonar, y establece la existencia de algún Dios, o hace observaciones relativas a los asuntos humanos; pero de pronto me asombra advertir que, en vez de las habituales cópulas para las proposiciones, *es* y *no es*, no hay una sola proposición cuya conjunción no sea un *debería* o un *no debería*.

» Este cambio es imperceptible; no obstante es de una importancia decisiva. Porque, puesto que el *debería* o el no *debería* expresan una nueva relación de afirmación, es necesario observarla y explicarla; y al mismo tiempo es necesario dar razón de algo que parecía completamente inconcebible: cómo es posible que esta nueva relación pueda deducirse de las otras, que son completamente distintas de ella.»

En este célebre pasaje de su *Tratado sobre la naturaleza humana*, el filósofo escocés David Hume brinda, con su habitual estilo lacónico, la clásica formulación de lo que desde entonces ha sido una de las cuestiones centrales de la filosofía moral. ¿Cómo es posible pasar de una proposición *descriptiva* acerca de cómo son las cosas en el mundo (juicios de hecho) a una proposición *prescriptiva* que nos dice qué deberíamos hacer (juicios de valor)? O, dicho brevemente, ¿cómo podemos derivar un « debe» de un « es» ? Evidentemente, Hume piensa que no podemos hacerlo y muchos otros pensadores comparten su criterio y consideran que « la guillotina de Hume» (o, dicho de forma más prosaica, « la ley de Hume») dividió el mundo entre hechos y valores de un modo decisivo.

Cronología

c. 440 a. C.	1739
La carne de un hombre...	La guillotina de Hume La teoría del abucheo y el hurra

La falacia naturalista La ley de Hume se suele confundir con otra idea relacionada pero distinta que planteó el filósofo inglés G. E. Moore en *Principios éticos* (1903). Moore acusaba a los primeros filósofos de incurrir en la llamada « falacia naturalista» que implica *identificar* los conceptos éticos con los conceptos naturales; así, por ejemplo, se considera que « bueno» significa lo *mismo que* (por ejemplo) « placentero» . Pero, de acuerdo con Moore, la pregunta sobre si lo placentero es también bueno sigue abierta —no es un asunto vacuo— de modo que la identificación sería falsa.

La posición de Moore (que ha sido menos influyente que la supuesta falacia que el autor identificó) consistía en que términos éticos como «bueno» son propiedades no naturales, es decir, son propiedades simples e inanalizables, sólo accesibles por medio de un sentido moral especial que se conoce con el nombre de «intuición».

Para complicar aún más las cosas, la expresión «falacia del naturalismo» también se usa en ocasiones para señalar un error completamente distinto —más propio de publicistas—, que consiste en pretender que el hecho de que algo sea natural (o artificial) es razón suficiente para suponer que también es bueno (o malo). Los pacientes de patógenos naturales que siguen tratamientos con fármacos sintéticos podrían atestiguar cabalmente hasta qué punto es falaz este argumento.

«**Tal vez el aspecto más simple y más importante de la ética sea estrictamente lógico. Me refiero a la imposibilidad de derivar normas éticas que no sean tautológicas … de juicios de hecho.**»

Karl Popper, 1948

1781 El imperativo categórico

1785 Los medios y los fines

-ismos éticos

La ética o la filosofía moral suele dividirse en tres grandes ámbitos. En el sentido más general, la **meta-ética** investiga el origen o las bases de la moralidad, incluyendo cuestiones como si es esencialmente objetiva o subjetiva por naturaleza. La **ética normativa** se centra en las normas éticas en las que se basa la conducta moral; así, el utilitarismo, por ejemplo, es un sistema normativo basado en la norma de la «utilidad». Por último, en el nivel más **concreto, la** ética aplicada traslada la teoría filosófica a los asuntos prácticos, como el aborto, la eutanasia, la guerra justa y el trato a los animales. Los filósofos han adoptado una serie de posiciones sobre todas estas cuestiones; y de ahí proceden los distintos -ismos. A continuación, ofrecemos un resumen esquemático de las posiciones éticas más comunes.

• El absolutismo

Sostiene que determinadas acciones son correctas o incorrectas en cualquier circunstancia.

• El consecuencialismo

Sostiene que la corrección o incorrección de las acciones puede establecerse limpiamente por recurso a su eficacia para conseguir determinados fines deseables o estados de hecho. El sistema consecuencialista más célebre es el **utilitarismo** (véase el capítulo 19).

• El deontologismo

Considera que determinadas acciones son intrínsecamente correctas o erróneas, con independencia de sus consecuencias; atribuye una importancia decisiva a las intenciones del agente y a las nociones de deberes y derechos. La **ética kantiana** es el sistema deontológico más importante (véase la página 76).

Los valores en un mundo desprovisto de valores El problema que Hume puso de relieve se debe en parte a dos convicciones fuertes en conflicto, de las que muchos de nosotros participamos simultáneamente. Por una parte creemos que vivimos en un mundo físico que, en principio, puede explicarse completamente por medio de leyes que la ciencia es capaz de descubrir; un mundo de hechos objetivos en el que los valores están excluidos. Pero, por otra parte, sentimos que al hacer juicios morales, por ejemplo cuando condenamos el genocidio, estamos afirmando algo cierto acerca del mundo; es algo que podemos saber y que debería ser verdad en cualquier caso, con independencia de los sentimientos que nos inspire. Pero estas dos concepciones parecen incompatibles si aceptamos la ley de Hume; y si no podemos fundamentar nuestros juicios morales en el mundo desprovisto de valores que describe la ciencia, aparentemente nos vemos forzados a volver a nuestros sentimientos y preferencias, y obligados a buscar en nuestro interior los orígenes de nuestros sentimientos morales. El propio Hume era consciente de la importancia de esta observación, y creía que si se le prestara la debida atención « todos los sistemas de moralidad habituales» se alterarían. La brecha lógicamente infranqueable entre hechos y valores, y que Hume parece haber abierto, pone en duda el estatuto verdadero de las pretensiones éticas y, así, subyace en el centro mismo de la filosofía moral.

• El naturalismo

Sostiene que los conceptos éticos pueden explicarse o analizarse estrictamente como « hechos de la naturaleza» que puede descubrir la ciencia, y suele tratarse de hechos de la naturaleza humana como el placer.

• El anticognitivismo

Considera que la moralidad no es un asunto de conocimiento, porque la preocupación de la moralidad no debe consistir en *los hechos*; por el contrario, un juicio moral expresa las actitudes, las emociones, etc., de la persona que actúa. Ejemplos de las posiciones anticognitivistas son el **emotivismo** y el **prescriptivismo** (véase el capítulo 16).

• El objetivismo

Sostiene que los valores morales y las propiedades forman parte del «mobiliario (o de la fábrica) del universo», y existen independientemente de los seres humanos capaces de aprehenderlos; las afirmaciones éticas no son subjetivas o relativas a nada, y deben ser ciertas o falsas de acuerdo con su mayor o menor adecuación al modo en que el mundo es. El objetivismo afirma que los conceptos éticos son metafísicamente reales y, en consecuencia, coincide en muchos aspectos con el **realismo** moral.

• El subjetivismo

Considera que el valor no se basa en la realidad externa sino en nuestras creencias acerca de la realidad o en nuestras reacciones emocionales a la misma. Esta posición coincide en lo fundamental con la del anticognitivismo (mencionada antes). En el caso de la posición cognitivista, el subjetivista sostiene que existen hechos éticos pero niega que sean objetivamente verdaderos o falsos; un ejemplo de esta forma de subjetivismo es el **relativismo** (véase el capítulo 14).

La idea en síntesis: la brecha entre el ser y el deber

13 La carne de un hombre…

«Cuando Darío era rey de Persia, hizo llamar a los griegos que había en su corte y les preguntó a cambio de qué estarían dispuestos a comer los cadáveres de sus padres. Los griegos contestaron que no lo harían por nada del mundo. Más tarde, en presencia de los griegos, y a través de un intérprete para que pudieran entender lo que se decía, Darío preguntó a unos indios de la tribu de los caladas, que de hecho se habían comido los cadáveres de sus padres, a cambio de qué estarían dispuestos a quemar los cadáveres [como era costumbre entre los griegos]. Prorrumpieron en un grito de horror y le prohibieron hablar de algo tan espantoso.»

¿Quién tiene razón, los griegos o los indios caladas? Inevitablemente, la perspectiva de comernos a nuestros padres nos hace palidecer, pero lo mismo les ocurre a los indios ante la perspectiva de quemar a los suyos. A fin de cuentas, seguramente estaríamos de acuerdo con Heródoto, el historiador griego que narra este relato, cuando cita con aprobación al poeta Píndaro: « La costumbre es el rey absoluto» . No se trata de que en un lado las cosas sean correctas y en el otro sean incorrectas: no existe una « respuesta correcta» . Cada comunidad tiene su propio código de costumbres y tradiciones; cada una se comporta correctamente de acuerdo con su propio código, y a él apelará cada comunidad para defender sus respectivos rituales funerarios.

Aquí lo moralmente correcto no parece ser absoluto, ni en un caso ni en el otro: es relativo a la cultura y las tradiciones de los grupos sociales en cuestión. Y naturalmente existen muchos más ejemplos de la diversidad cultural, tanto geográfica como histórica. En casos como éste el relativista argumenta que, en general, no existen verdades absolutas o universales: todas las valoraciones y los juicios morales deberían hacerse sólo en relación con las respectivas normas sociales de los grupos.

Cronología

440 a. C.
La carne de un hombre...

Vive la différence La propuesta del relativista consiste, en efecto, en abordar los juicios morales como si fueran juicios estéticos. En materia de gustos no suele ser adecuado hablar de error: *de gustibus non disputandum* (« sobre gustos no hay discusión posible»). Si dices que te gustan los tomates y a mí no me gustan, no tendremos inconveniente en diferir: algo es cierto o falso para ti, pero no para mí. En estos casos la verdad tiene que ver con la sinceridad: si digo sinceramente que me gusta algo, no puedo equivocarme: es verdad (para mí). Y siguiendo esta

analogía, si nosotros (como sociedad) aprobamos la pena de muerte, es moralmente correcta (para nosotros), y no hay nada en lo que podamos estar equivocados. Del mismo modo que no intentaríamos persuadir a la gente de que dejara de comer tomates ni los criticaríamos por hacerlo, en el caso de la moral la persuasión o la crítica serían inapropiadas. Es evidente que, de hecho, nuestra vida moral está llena de razones y de censura, y solemos adoptar posiciones fuertes en asuntos como la pena de muerte. Incluso es posible que terminemos discrepando con *nosotros mismos* al cabo del tiempo: puedo cambiar de parecer sobre alguna cuestión moral, y hasta las posiciones colectivas pueden cambiar, por ejemplo en un asunto como el de la esclavitud. El relativista consumado diría que algo era correcto para algunos pero no para otros, o correcto para mí (o para nosotros) en una época pero no en otra. Casos como la esclavitud, la ablación, el infanticidio legal, etc., pueden convertirse para el relativista en una píldora amarga de tragar.

«**¿Qué es la moralidad en una determinada época o lugar? Es lo que a la mayoría le gusta en ese momento y en ese lugar, y la inmoralidad es lo que le disgusta.**»

Alfred North Whitehead, 1941

El fracaso del relativismo para ofrecer una explicación seria de aspectos que son tan evidentemente característicos de nuestras vidas morales en la actualidad es considerado a menudo como un golpe decisivo contra la tesis, pero los relativistas pueden intentar sacar ventaja del mismo. Quizás, podrían argumentar, no deberíamos ser tan inquisidores y críticos con los otros. La lección de los griegos y los indios caladas es que debemos ser más tolerantes con los otros, tener una mentalidad más abierta, ser más sensibles a sus costumbres y a sus prácticas. Esta línea argumentativa ha llevado a muchos a asociar el relativismo con la tolerancia y la mentalidad abierta, y por contraste se representa a los no relativistas como individuos intolerantes a los que exaspera toda práctica distinta a las suyas. Se supone que el antirrelativismo, llevado al extremo, desemboca en el imperialismo cultural occidental que impone con arrogancia sus criterios a los demás y los arrolla. Pero esto es una caricatura: de hecho, no existe incompatibilidad alguna entre adoptar una posición tolerante en general y mantener que algunas personas o algunas culturas están equivocadas con respecto a algunos asuntos. Por lo demás, un inconveniente con el que topa el relativista es que ¡sólo el que no lo es puede abrazar la tolerancia y la sensibilidad cultural como virtudes universales! (véase el cuadro).

Círculos viciosos del relativismo

El relativismo fuerte o radical (la idea de que todos los juicios —morales o de cualquier otro tipo— son relativos) da lugar de inmediato a una serie de nudos.

¿El juicio según el cual todos los juicios son relativos es asimismo relativo? Efectivamente debería serlo, para evitar autocontradecirse; pero si es relativo ello significa que la pretensión de que todas mis reivindicaciones son absolutas es cierta *para mí*. Y este tipo de incoherencias termina intoxicando todo lo demás. Los relativistas no pueden afirmar que criticar las convenciones culturales de otras sociedades es siempre un error, pues *para mí* puede ser correcto hacerlo. Y no pueden defender que siempre sea correcto ser tolerante y de mentalidad abierta, pues para algún autócrata podría ser correcto eliminar cualquier signo de disentimiento. En general, los relativistas no pueden sostener la validez de su propia posición de un modo consistente y que evite la hipocresía. La naturaleza autorrefutadora del relativismo radical incipiente puede verse en Platón, quien muy pronto mostró las inconsistencias en la posición relativista adoptada por el sofista Protágoras (en un diálogo que lleva por título su nombre). La lección que extraemos es que la discusión racional depende de la posibilidad de compartir *algún* punto fundamental: debemos estar de acuerdo en *algo*, compartir una verdad común, para que la comunicación sea significativa. Pero precisamente este fundamento común es lo que niega el relativismo radical.

El conocimiento en perspectiva Lo absurdo del verdadero relativismo y los peligros de su adopción generalizada como un mantra político (véanse los cuadros siguientes) han implicado que algunos de los aciertos de las formas más moderadas de relativismo queden a veces relegados. La lección más importante del relativismo es que el conocimiento mismo está sujeto a la perspectiva: siempre miramos el mundo desde una determinada perspectiva o punto de vista; no existe una posición externa privilegiada desde la que observar el mundo « tal como es en realidad» o « como es en cualquier caso» . Esta posición se explica a menudo a partir de esquemas conceptuales o marcos: dicho llanamente, sólo podemos alcanzar una comprensión intelectual de la realidad desde el interior de nuestro propio marco conceptual, determinado a su vez por una compleja combinación de factores que incluyen nuestra cultura y nuestra historia. Pero que no podamos sustraernos o salir de nuestro particular esquema conceptual y

adoptar una visión objetiva de las cosas —« la visión de Dios» — no significa que no podamos conocer nada. Una perspectiva debe serlo de *algo*, y compartiendo y comparando nuestras distintas perspectivas podemos esperar compensar nuestras diversas creencias, y alcanzar una representación del mundo más completa, más abarcadora, más « estereoscópica» . Esta imagen benigna indica que el progreso hacia la comprensión sólo puede lograrse a través de la colaboración, de la comunicación y del intercambio de ideas y de puntos de vista: un legado muy positivo del relativismo.

¿Vale todo?

« Actualmente la tarea de educar se ve dificultada por un obstáculo particularmente insidioso: la absoluta presencia en nuestra sociedad y en nuestra cultura de un relativismo que no reconoce nada como definitivo y, así presenta como último y único criterio el yo y sus deseos. Y bajo la apariencia de libertad se convierte en una prisión para cada cual, pues separa a la gente entre sí, y la deja encerrada en la contemplación de su propio "ego".»

Papa Benedicto XVI, junio de 2005

A lo largo de las últimas décadas, la idea del relativismo ha cobrado una relevancia social y política que ha extendido su significado original hasta el límite del paroxismo.

De la idea de que no existen verdades absolutas —« todo es relativo» — se ha inferido que todo es igualmente válido y, así, que « todo vale» . Distintas fuerzas reaccionarias, entre ellas sectores de la Iglesia católica, creen que el hecho de que algunos de estos razonamientos hayan cobrado protagonismo explica el libertinaje moral (y especialmente sexual) y la desintegración social, que atribuyen a las fuerzas relativistas mundiales. Por otro lado, a quienes suscriben el liberalismo les complace tanto el relativismo que ni siquiera han considerado en serio su lógica, y han abrazado el « todo vale» como su mantra político. De modo que los bandos opuestos se han agrupado, unos en las filas del placer y los otros en las del horror, en torno a un amedrentado relativismo.

La idea en síntesis: ¿es todo relativo?

14 La teoría del mandato divino

Establecer qué es lo correcto y lo erróneo, lo bueno y lo malo, la virtud y el vicio, es el típico problema del que sólo cabe esperar que nos quite el sueño: el aborto, la eutanasia, los derechos humanos, el trato a los animales, la investigación con células madre… una lista interminable de asuntos peliagudos y comprometedores. Más que ningún otro ámbito, la ética inspira la sensación de andar en un campo minado: es un terreno traicionero donde prevés perder pie en cualquier momento, y donde, además, los tropiezos pueden costar un alto precio.

Y aun así, paradójicamente, para mucha gente dedicarse a moralizar resulta aparentemente algo así como un paseo por el parque. Para la mentalidad de millones de personas, la moralidad se halla unida de forma inextricable a la religión: esto o lo otro es correcto o incorrecto por el simple hecho de que Dios (o un dios) ha establecido que así sea: lo bueno es bueno y lo malo es malo porque así lo dice Dios.

En cada una de las tres « religiones del Libro» —el judaismo, el cristianismo y el islam—, el sistema de la moralidad se basa en un « mandato divino» : lo propio de Dios es ordenar y lo propio de los hombres obedecer; Dios impone a sus devotos una serie de mandamientos morales; el comportamiento virtuoso exige obediencia, mientras que la desobediencia es pecado. Seguramente tal código de normas éticas, garantizado por la mano de Dios, permite disipar las preocupaciones que acucian a las explicaciones subjetivas de la moralidad: la desagradable sospecha de que vamos estableciendo las normas sobre la marcha.

Cronología

c. 375 a. C.	1670
Teoria del mandato divino	Fe y razón

«Ninguna moralidad debería basarse en la autoridad, ni siquiera si esa autoridad fuera divina.»

A. J. Ayer, 1968

El dilema de Eutifrón Naturalmente, sin Dios la teoría del mandato divino se hunde de inmediato (véase el cuadro de la página siguiente), pero incluso aceptando que Dios exista, persiste un buen número de problemas serios que amenazan la teoría. Posiblemente el más grave de ellos sea el que se conoce como el dilema de Eutifrón, presentado por primera vez por Platón hace unos 2 400 años, en su diálogo titulado *Eutifrón*. Sócrates (el portavoz de Platón en los

diálogos) hace participar en una discusión sobre la naturaleza de la piedad a un joven llamado Eutifrón. Los dos están de acuerdo en que la piedad es « todo aquello que aman los dioses» , pero inmediatamente Sócrates plantea una pregunta crucial: ¿son piadosas las cosas porque los dioses las aman, o las aman porque son piadosas? Este dilema (expresado por lo general en un contexto monoteísta) pone en apuros a la teoría del mandato divino.

Comprender los mandados divinos

Si dejamos a un lado el dilema de Eutifrón, existe otra dificultad que debe afrontar quien basa la moralidad en el mandato divino: los distintos textos religiosos que constituyen el principal medio de manifestación de la voluntad de Dios a los humanos contienen muchos mensajes encontrados y/o inasumibles. Basta tomar un ejemplo conocido de la Biblia, en el Levítico (20:13), donde leemos: « Si un hombre yace junto a un hombre como se hace con una mujer, los dos han cometido una abominación; morirán sin remedio; su sangre caerá sobre ellos» . Si la Biblia es la palabra de Dios y la palabra de Dios determina lo que es moral, la ejecución de los varones homosexuales está moralmente autorizada. Pero en la actualidad la mayoría consideraría semejante idea como moralmente repugnante, y en cualquier caso resulta inconsistente con los mandamientos que encontramos en otros pasajes de la Biblia (el más obvio, el de no matar). Es evidente que para el teórico del mandato divino es todo un reto usar los conocidos valores de Dios para construir un sistema moral aceptable, de un modo general e internamente coherente.

1739	1958
La teoría del abucheo y el hurra La guillotina de Hume	Bajo la llamada del deber

¿Qué hacer?

El mayor peligro al que debe hacer frente la teoría del mandato divino es la pérdida de su ordenante divino: es preciso que estemos perfectamente persuadidos de los distintos argumentos que se presentan para probar la existencia de Dios, y no podemos disfrutar del beneficio de la fe (véase el capítulo 44).

Sin inmutarse, algunos de los abogados de la teoría han hecho de la necesidad virtud al utilizar este peligro como prueba de la existencia de

Dios:

1. Existe algo como la moralidad: tenemos un código de leyes o mandatos éticos.

2. Dios es el único candidato para el papel de legislador/gobernante, de modo que

3. Dios debe existir.

Sin embargo, es poco probable que esta argumentación sirva ante un oponente real. La primera premisa, que implica que la moralidad es esencialmente algo que existe con independencia de los humanos, da por descontada una de las cuestiones de fondo más fundamentales. E incluso concediendo que la moralidad existiera independientemente de nosotros, la segunda premisa tiene que resistir la embestida del ataque de Eutifrón.

Así pues ¿lo bueno es bueno porque Dios lo establece, o Dios lo establece porque es bueno? Ninguna de las dos alternativas le resulta aceptable al teórico del mandato divino. Consideremos en primer lugar la primera parte: matar (por poner un ejemplo) se considera incorrecto porque Dios lo establece, pero las cosas podrían ser de otro modo. Dios podría haber establecido que matar es correcto o incluso obligatorio, y lo habría sido (sólo porque Dios lo habría querido así). De acuerdo con esta lectura, la observancia religiosa aporta poco más que la ciega obediencia a una autoridad arbitraria. Pero ¿acaso la otra alternativa funciona mejor? No: si Dios establece lo que es bueno porque es bueno, es evidente que la bondad es independiente de Dios. En el mejor de los casos, el papel de Dios es el del mensajero moral, que transmite las prescripciones éticas pero no es su fuente. De modo que podríamos ir directamente a la fuente y matar al mensajero. Por lo menos como legislador el papel de Dios es redundante. De modo que, en lo que se refiere a la moralidad, Dios es o bien arbitrario o bien irrelevante. No es una alternativa fácil para quienes buscan poner a Dios como garante o autoridad de su ética.

«¿Aman los dioses al piadoso porque es piadoso, o acaso es piadoso porque lo aman?»

Platón, *c*. 375 a. C.

Un contraataque habitual al dilema de Eutifrón es insistir en que «Dios es bueno», y en consecuencia, no establecería algo malo. Pero esta línea de ataque corre el riesgo de incurrir en circularidad o en incoherencia. Si «bueno» significa «establecido por Dios», «Dios es bueno» resultaría una expresión prácticamente carente de sentido: algo así como «Dios obedece sus propios mandatos».

Tal vez resulte más fructífero considerar que la frase significa «Dios es (idéntico) a la bondad» y, en consecuencia, que sus mandatos serán inevitablemente buenos. Pero si la divinidad y la bondad son una misma cosa, «Dios es bueno» es una expresión del todo vacua: no arroja luz y es circular (un ejemplo, tal vez, de la afición de Dios a las sendas inescrutables).

La idea en síntesis: porque Dios lo establece

15 La teoría del abucheo y del hurra

«Y Moisés permaneció allí con el Señor cuarenta días y cuarenta noches; no comió nada de pan ni bebió nada de agua. Y escribió sobre las tablas las palabras de la alianza, es decir, los diez mandamientos.
—¡Hurra! ¡Por no tener a ningún otro Dios ante mí!
—¡Abajo los que te erijan ídolos!
[cinco abucheos y dos hurras se sucedieron; entonces…]
—¡Abajo los que codicien a la mujer del vecino, o a sus criados, o a sus criadas, o a su buey, o a su asno, o a cualquier cosa que sea de su vecino!»

Así habló el Señor, de acuerdo con el emotivismo, o la teoría ética del «abucheo/hurra». Así presentado, el emotivismo no parece una tentativa seria para explicar la fuerza de las afirmaciones éticas (y esta impresión parece reforzarla el carácter desenfadado del título que recibe la teoría). De hecho, sin embargo, el emotivismo es una teoría muy influyente con un pedigrí considerable, y su motivación es la profunda preocupación por lo que podría resultar una comprensión de nuestras vidas morales más acorde con el sentido común.

La hora del subjetivismo Existen distintos tipos de hechos en el mundo que son objetivamente ciertos: son hechos cuya verdad no depende de nosotros. Algunos de ellos son científicos, describen acontecimientos físicos, procesos o relaciones; otros son morales y describen cosas del mundo que son correctas o incorrectas, buenas o malas. Este esquema tal vez podría interpelar al sentido común, pero para muchos filósofos está visto que resulta muy poco atractivo.

Cronología

c. 375 a. C.	*c.* 30 d. C.
La teoría del mandato divino	La regla de oro

La razón, esclava de las pasiones

La principal inspiración de las formas modernas de subjetivismo moral es el filósofo escocés David Hume. Su famoso alegato de una explicación subjetivista de la moralidad aparece en su *Tratado de la naturaleza humana*:

«Consideremos cualquier acción reputada de viciosa: el asesinato premeditado, por ejemplo. Examinémosla bajo distintas luces, y veamos si es posible encontrar el hecho, o la existencia real, de eso a lo que llamamos vicio. Por más consideraciones distintas que hagamos, sólo

encontraremos algunas pasiones, motivos, voliciones y pensamientos. No hay más hechos en el asunto. El vicio se nos escapa completamente tan pronto como examinamos el objeto. Nunca puedes dar con él, hasta que tu reflexión se vuelve hacia tu propio interior y encuentras un sentimiento de desaprobación, que surge de ti, con respecto a ese acto. Esto es un hecho; pero éste es un objeto del sentimiento no de la razón. Reside en ti, no en el objeto».

De acuerdo con la explicación que Hume da de la acción moral, todos los humanos actúan movidos por un «sentido moral» o «simpatía» natural, que esencialmente consiste en una capacidad para compartir los sentimientos de felicidad o de desdicha de los otros; y este sentimiento, más que la razón, es el que proporciona el motivo a nuestras acciones morales. La razón es esencial para comprender las consecuencias de nuestras acciones y para planificar racionalmente cómo alcanzar nuestros objetivos morales, pero en sí misma es inútil e incapaz de proporcionar cualquier impulso a la acción: tal como dijo el propio Hume en una famosa frase, «la razón es, y debería ser sólo, la esclava de las pasiones».

«**No existe nada bueno ni malo, es el pensamiento el que lo hace aparecer así.**»

William Shakespeare, *c.* 1600

Consideremos un hecho moral: matar es incorrecto. Podemos describir el acto de matar de un modo muy detallado, mencionando todo tipo de hechos físicos y psicológicos para explicar cómo y por qué se llevó a cabo. Pero ¿qué propiedad o cualidad adicional podemos añadir a la descripción cuando lo consideramos incorrecto? Fundamentalmente estamos afirmando que matar es el tipo de cosa que no *deberíamos* hacer (que, entre todas las demás cosas que sinceramente podemos decir del matar, también tiene una propiedad intrínseca que consiste en «no-deber-ser-realizable»). Perplejos ante la completa excentricidad de tener que encontrar una propiedad semejante en el mundo (el mundo supuestamente desprovisto de valores que describe la ciencia; véase la página 136), muchos filósofos proponen reemplazar la noción de las propiedades morales objetivas que existen en el mundo por una respuesta subjetiva a las cosas del mundo.

1739 La teoría del abucheo y del hurra

1974 ¿Cómo es ser un murciélago?

El prescriptivismo

La crítica más habitual al emotivismo consiste en señalar que no consigue captar la lógica del discurso ético (la forma característica del razonamiento y del argumento racional que subyace a ese discurso). Conseguirlo constituye uno de los principales propósitos de una teoría que rivaliza con el subjetivismo, y que se conoce con el nombre de prescriptivismo, estrechamente vinculada a la figura del filósofo inglés R. M. Hare. El punto de partida de esta teoría es la idea de que los términos morales poseen un elemento prescriptivo —nos indican qué hacer o cómo comportarnos—, a partir de lo cual propone que la esencia de los términos morales consiste en orientar la acción; decir que matar es erróneo equivale a asumir un mandato: «No matarás». De acuerdo con la explicación de Hare, el elemento que distingue los juicios éticos de otros tipos de mandato es que los primeros son «universalizables»: si establezco un mandato moral, me comprometo a sostener que tal mandato debería ser obedecido por cualquiera (incluido yo mismo) en ocasiones similares (i. e. debo cumplir la regla de oro; véase la página 80). La discrepancia moral, de acuerdo con el prescriptivismo, es análoga a dar órdenes contradictorias; la inconsistencia y la indecisión se explican porque existen distintos mandatos, algunos de los cuales no pueden ser obedecidos simultáneamente. En este sentido, el prescriptivismo deja más espacio para la discrepancia y el debate que el emotivismo, aunque hay quien sigue preguntándose si, en verdad, refleja toda la complejidad del diálogo moral.

De la descripción a la expresión De acuerdo con un punto de vista subjetivista ingenuo, los juicios morales simplemente describen o informan de nuestros sentimientos acerca del modo en que las cosas son en el mundo. Así, cuando digo «El asesinato no está bien», sólo afirmo mi desaprobación (o tal vez la de mi comunidad). Pero esto es demasiado simple. Si digo «el asesinato está bien» y asumo que ésa es una descripción adecuada de mis sentimientos, entonces también será verdadero. Y en ese caso la discrepancia moral sería aparentemente imposible. De modo que se requiere algo más sofisticado.

El emotivismo (o expresivismo) —la teoría del abucheo y del hurra— constituye una forma más sutil de subjetivismo, al sugerir que los juicios morales no son descripciones o declaraciones de nuestros sentimientos acerca del mundo sino

«**No se me ocurre cómo podrían refutarse los argumentos sobre la subjetividad de los valores éticos, pero al mismo**

expresiones de esos sentimientos. De modo que, cuando hacemos un juicio moral, estamos expresando una respuesta emocional: nuestra aprobación («¡hurra!») o nuestra desaprobación (el abucheo) con respecto a algo del mundo; «es bueno decir la verdad» es una expresión de nuestra aprobación («¡hurta por decir la verdad!»). El gran problema de los emotivistas es aproximar su teoría de algún modo a la manera en que pensamos realmente sobre nuestro discurso moral, y en que actuamos de acuerdo con ella. Este discurso presupone un mundo externo de valores objetivos: reflexionamos y discutimos sobre problemas morales; apelamos a los hechos morales (y a otros) para resolverlos; realizamos juicios éticos que pueden ser verdaderos o falsos; y existen verdades morales que podemos llegar a conocer. Pero de acuerdo con los emotivistas, no existe el conocimiento ético (no estamos afirmando nada en absoluto, sino que estamos expresando nuestros sentimientos, y tales expresiones no pueden ser verdaderas o falsas). El emotivismo admite que la deliberación y el desacuerdo son posibles a partir de nuestras creencias de fondo y del contexto de nuestras acciones, pero resulta difícil darles consistencia a partir de algo como nuestra concepción habitual del debate moral. Las conexiones lógicas entre los propios juicios morales parecen estar ausentes, y el razonamiento moral es aparentemente poco más que un ejercicio de retórica (la moralidad como discurso publicitario, tal como se ha insinuado cáusticamente). La respuesta invariable a este problema consiste simplemente en resignarse: sí, responde el emotivista, la teoría no concuerda con nuestros supuestos habituales, pero ello ocurre porque esos supuestos, y no la teoría, son erróneos. De acuerdo con este presunto «error teórico», nuestro discurso ético normal simplemente está equivocado, porque se basa en hechos morales objetivos que no existen realmente. Son muchas las tentativas que han perseguido aproximar la representación emotivista a nuestro discurso ético supuestamente realista, pero a pesar de todo la distancia sigue siendo grande, de modo que se han propuesto otras alternativas. Posiblemente la más importante de ellas sea el prescriptivismo (véase el cuadro anterior).

tiempo me resulta imposible pensar que lo único malo de la crueldad gratuita es que no me gusta.»

Bertrand Russell, 1960

La idea en síntesis: expresar juicios morales

16 Medios y fines

«El señor Quelch no tenía del todo claro si los tiburones tienen labios y, en caso de tenerlos, si pueden lamérselos; pero lo que sí tenía claro es que si tenían labios y se los podían lamer, eso debía ser exactamente lo que estaban haciendo entonces. En aquel momento el globo se precipitaba hacia el mar, y pudo ver con claridad las múltiples aletas de los comensales reunidos cortando amenazadoramente el agua…

» … el señor Quelch sabía que en los dos minutos siguientes él mismo y el puré de sus cachorros serían el cebo de los tiburones, a menos que consiguieran echar lastre. Pero ya había arrojado todo fuera de la cesta: lo único que quedaba era él y los seis chicos. Estaba claro que sólo Guillermito era un paquete de suficiente peso para resolver la papeleta. Un hueso duro de roer incluso para Guillermito el del gran apetito, pero no quedaba más remedio…

—Ay, cielos… ay, realmente, amigos míos… mirad que si me ponéis un dedo encima… ¡Ayayay!»

Un avión comercial en el que viajan 120 pasajeros vuela a toda velocidad fuera de control en dirección a una zona muy poblada. No hay tiempo de evacuar la zona y el impacto del avión matará con toda seguridad a cientos de personas. La única opción para evitarlo es hacer caer al avión. ¿Habría que hacerlo?

Supongamos que la evaluación de la situación que hace el señor Quelch sea perfectamente justa. Sólo existen dos opciones: los seis chicos, incluido Guillermito, y el propio Quelch, caen al mar y los tiburones los hacen pedazos; o sólo se arroja al mar y se sacrifica a Guillermito. Aunque ser arrojado al mar es desagradable, para Guillermito existe poca diferencia, pues morirá en cualquier caso, mientras que echando a Guillermito del globo Quelch puede salvarse y salvar a los otros cinco chicos. ¿Es correcto entonces sacrificar a Guillermito? ¿Acaso el fin, salvar varias vidas inocentes, justifica los medios, sacrificar una vida inocente?

Cronología

c. 1260	1739	1781
Actos y omisiones Guerra justa	La guillotina de Hume	El imperativo categórico

Un dilema ético Las decisiones donde están en juego la vida y la muerte no son, evidentemente, simples temas de ficción. En la vida real, a veces

«El fin puede justificar los medios en la medida en que exista algo que justifique el

nos encontramos en situaciones en las que es necesario permitir que muera uno o algunos inocentes, o en los casos extremos incluso permitir que sean asesinados, para que se salven otras muchas vidas inocentes. Se trata de casos que ponen radicalmente a prueba nuestras intuiciones, casos que nos dividen repentinamente entre dos opciones, y a veces resulta tan difícil escoger que nos desgarran.

fin.»

Leon Trotski, 1936

Las diversas aproximaciones de los filósofos para intentar explicar dilemas de este tipo reflejan esta incertidumbre fundamental. Se considera que las distintas teorías que se han propuesto se inscriben en uno u otro lado de la línea divisoria fundamental en ética: la línea que separa las teorías basadas en el deber (deontológicas) de las teorías basadas en las consecuencias (consecuencialistas).

Consecuencialismo y deontología Un modo de poner de relieve las diferencias entre el consecuencialismo y la deontología es el punto de vista de los medios y los fines. El consecuencialismo propone que saber si una acción es buena o mala debería establecerse estrictamente en razón de sus consecuencias; una acción se considera sólo como un medio para alcanzar un fin deseable, y su acierto o error depende de cómo de efectivo resulte para alcanzar ese fin. El fin en sí mismo es algún estado de hecho (tal como la felicidad) que resulta de, o es consecuencia de, las distintas acciones que contribuyen a alcanzarlo.

Dos hermanos siameses morirán con toda seguridad en unos pocos meses a menos que se sometan a una operación quirúrgica para separarlos. La necesaria operación ofrece un pronóstico excelente para uno de los gemelos, que tendrá una vida razonablemente saludable y plena, pero para ello es necesario que el otro gemelo muera. ¿Harías la operación? (¿La harías incluso si los padres no dieran su consentimiento?)

Al escoger entre distintos modos disponibles de actuar, los consecuencialistas sólo sopesarán las consecuencias buenas y malas en cada caso, y tomarán sus decisiones a partir de esta evaluación. En el caso de Guillermito, por ejemplo, se inclinarán a juzgar que el buen resultado en términos de vidas inocentes salvadas justifica el sacrificio de una vida.

1785	1954	1974
Medios y fines	Terreno resbaladizo	La máquina de la experiencia

En cambio, en un sistema deontológico las acciones no son consideradas como meros medios para alcanzar un fin sino como acciones correctas o incorrectas

por sí mismas. Es necesario que las acciones estén provistas de un valor intrínseco por derecho propio, y no sólo de un valor instrumental que contribuya a la consecución de un fin deseable. Por ejemplo, el deontólogo dictaminaría que matar a inocentes es intrínsecamente malo: arrojar a Guillermito por la borda es incorrecto en sí mismo, y no existe ninguna consecuencia bondadosa que pueda justificar tal acción.

El caso de Guillermito puede parecer exagerado, pero a veces la vida real nos plantea dilemas complicados parecidos. Todos los casos que se han recogido en este capítulo son similares —al menos en lo que se refiere a las dudas morales que suscitan— a acontecimientos que ocurren de hecho y que, sin duda, seguirán ocurriendo una y otra vez.

La teoría consecuencialista más conocida es el utilitarismo (véase el capítulo 18); y el sistema deontológico más influyente es el que propuso Kant (véase el capítulo 19).

El fin justifica los medios

En un sentido trivial, un medio sólo puede justificarse por un fin, puesto que el medio es por definición una manera de alcanzar el fin; de modo que un medio está justificado (i. e. validado como medio) por el hecho exclusivo de contribuir al logro del fin. Pero pueden surgir problemas —y la máxima resultar entonces siniestra— cuando se escoge un fin inadecuado y la elección se hace a la luz de una ideología o de un dogma. Si, por ejemplo, un ideólogo político, o un fanático religioso, establece la importancia de un determinado fin para excluir a los demás, sus seguidores pueden llegar fácilmente a la conclusión de que desde el punto de vista moral es aceptable usar cualquier medio disponible para alcanzar ese fin.

Un paciente A terminal va a morir con toda seguridad en una semana. Su corazón y sus riñones son del todo compatibles para los pacientes B y C, que morirán ineluctablemente antes que él si no se les realiza el trasplante que necesitan, con el cual, no obstante,

tienen un buen pronóstico de recuperación. No existen otros donantes compatibles. ¿Matarías al paciente A (¿con su consentimiento, sin su consentimiento?) para salvar a los pacientes B y C?

Tú, junto con otros pasajeros y la tripulación de un pequeño aeroplano, sobrevivís a un accidente en medio de unas montañas desoladas. No disponéis de ningún tipo de comida, no hay modo de huir a pie y no hay esperanzas de que un equipo de rescate os encuentre hasta dentro de algunas semanas, y para entonces todos habréis muerto de hambre. La carne de un pasajero permitiría subsistir a los demás hasta que llegara la ayuda. ¿Matarías a uno de tus compañeros? (¿Cómo lo escogerías?)

Un oficial de la Gestapo ha atrapado a 10 niños y pretende matarlos a menos que tú reveles la identidad y el paradero de un espía. Casualmente ni siquiera tenías noticia de la existencia de un espía, y mucho menos de su identidad, pero sabes perfectamente no sólo que el oficial no te creerá si alegas ignorancia sino que además cumplirá su amenaza. ¿Dirías algún nombre —cualquiera— para salvar a los niños? (¿Y cómo decidirías qué nombre dar?)

La idea en síntesis: la opción menos mala

17 La máquina de la experiencia

«Supongamos que existiera una máquina de la experiencia que pudiera suministrarte alguna experiencia que desearas. Unos neuropsicólogos maravillosos estimularían tu cerebro de modo que pensarías y sentirías que estás escribiendo una novela genial, o haciendo amistad con alguien, o leyendo un libro interesante. Durante todo el tiempo estarías flotando en un gran recipiente con electrodos conectados a tu cabeza. ¿Te conectarías a esta máquina para siempre, programando anticipadamente los deseos de toda tu vida? … Naturalmente mientras estuvieras en el recipiente no sabrías que lo estás; pensarías que todo está ocurriendo de veras … ¿Te conectarías? ¿Qué más nos importa, aparte de cómo sentimos interiormente nuestras vidas?»

«Entre el sufrimiento y la nada, me quedaría con el sufrimiento.»

William Faulkner, 1939

El creador de este experimento mental en el año 1974, el filósofo norteamericano Robert Nozick, piensa que las respuestas a sus últimas preguntas son, respectivamente: «No» y «Mucho». A simple vista, la máquina de la experiencia se parece mucho al cerebro en la cubeta de Putnam (véase la página 8). Los dos experimentos describen realidades virtuales en las que se simula un mundo que resulta completamente indiscernible, por lo menos desde dentro, de la vida real. Pero mientras que Putnam se interesa por la situación del cerebro en la cubeta, y por las implicaciones que tiene para el establecimiento de los límites del escepticismo, la principal preocupación de Nozick es la situación de las personas antes de conectarse a la máquina: ¿escogerían una vida conectada a la máquina? y, si lo hicieran, ¿qué nos revelaría su elección?

Cronología

c. 1260	1739
Actos y omisiones	La guillotina de Hume

«La naturaleza ha puesto a la humanidad bajo el gobierno de dos amos soberanos, el dolor y el placer. Sólo ellos pueden indicarnos qué hacer.»

Jeremy Bentham, 1785

Se trata de escoger entre una vida simulada de ilimitado placer, en la que cada ambición y cada deseo se colmaran; y una vida real marcada por todas las previsibles frustraciones y decepciones, la mezcla habitual de sucesos incompletos y de sueños incumplidos. Pero a pesar del evidente atractivo de la vida conectada a la experiencia de la máquina, la mayoría de la gente, cree Nozick, no escogería que la conectaran. La *realidad* de la vida es importante: queremos hacer determinadas cosas, no sólo experimentar el placer de hacerlas. Pero si el

placer fuera lo único que determinara nuestro bienestar, si fuera el único elemento constitutivo de la vida buena, seguramente no escogeríamos la vida real, puesto que es mucho mayor el placer que obtendríamos conectados a la máquina de la experiencia. A partir de esta idea, Nozick infiere que, además del placer, existen otras cosas que consideramos intrínsecamente valiosas.

El utilitarismo clásico Esta conclusión da al traste con cualquier teoría ética hedonista (basada en el placer), y en particular con el utilitarismo, al menos en la formulación clásica que dio su fundador, Jeremy Bentham, en el siglo XVIII. El utilitarismo es la concepción para la cual las acciones deben juzgarse como buenas o malas en atención a su capacidad para incrementar o reducir el bienestar humano o la « utilidad» . Desde Bentham se han propuesto múltiples interpretaciones de la utilidad, pero para él consistía en la felicidad y el placer humanos, y su teoría de las acciones correctas se resume en ocasiones como el fomento de « la mayor felicidad del mayor número posible» .

Al utilitarismo no le asustan las conclusiones morales que se oponen a nuestras intuiciones habituales (véase la página 68). Para Bentham, una de sus principales virtudes consistía en que proporcionaba una base racional y científica a las decisiones sociales y morales, a diferencia de lo que ocurría con las caóticas e incoherentes intuiciones en las que se basaban los llamados derechos naturales y la ley natural. Con el propósito de brindar este fundamento racional, Bentham propuso un « cálculo de la felicidad» , de acuerdo con el cual las distintas cantidades de placer y dolor que producía cada acción podían medirse y compararse; la acción correcta en una ocasión dada podía determinarse entonces mediante un simple proceso de suma y resta. Así, para Bentham los distintos placeres difieren únicamente en cuanto a la duración y la intensidad, no en cuanto a su calidad; se trata de una concepción del placer considerablemente monolítica que parece vulnerable a las implicaciones de la máquina de la experiencia de Nozick. Dada su naturaleza inflexible, cabe sospechar que Bentham hubiera estado encantado de machacar la idea que el experimento mental de Nozick plantea. Sin embargo, a]. S. Mill, otro de los padres fundadores del utilitarismo, sí que le preocupaba limar algunas asperezas de la teoría.

Variedades del utilitarismo

El utilitarismo es, históricamente, la versión más relevante del consecuencialismo, es decir, de la concepción según la cual las acciones

deben juzgarse como correctas o incorrectas a la luz de sus consecuencias (véase la página 69). En el caso del utilitarismo, el valor de las acciones lo determina su contribución al bienestar o la «utilidad». En el utilitarismo clásico (hedonista) de Bentham y Mill, la utilidad es entendida como el placer humano, pero esta concepción se ha modificado y ampliado en distintos sentidos desde entonces. Las distintas versiones reconocen siempre que la felicidad humana depende no sólo del placer sino también de la satisfacción de una vasta gama de deseos y preferencias. Algunos teóricos también han propuesto ampliar el alcance del utilitarismo más allá del bienestar humano para abarcar otras formas de la vida sensible.

También existen diversas concepciones sobre cómo debe aplicarse el utilitarismo a las acciones. De acuerdo con el **utilitarismo de acción** o **directo**, cada acción debe evaluarse en atención a su contribución a la utilidad. En cambio, de acuerdo con el **utilitarismo normativo**, el desarrollo adecuado de la acción lo determinan distintas normas cuyo seguimiento general promoverá la persona inocente puede contribuir en determinadas ocasiones a salvar muchas vidas y, en consecuencia, incrementará la utilidad general, de manera que para el utilitarismo de acción éste sería un modo de actuar correcto. Sin embargo, como *regla*, matar a un inocente reduce la utilidad, de modo que el utilitarismo normativo sostiene que la misma acción es incorrecta, incluso aunque pudiera tener consecuencias beneficiosas en alguna ocasión particular. El **utilitarismo normativo** coincidiría así con nuestras convicciones comunes acerca de los asuntos morales, aunque los utilitaristas más recientes no se encomienden necesariamente a este principio: por distintas razones les parece incoherente cuando no objetable.

«Las acciones son correctas en relación con su capacidad para promover la felicidad, e incorrectas cuando tienden a producir lo contrario de la felicidad.»

J. S. Mill, 1859

Placeres elevados y bajos Los críticos contemporáneos se han apresurado a señalar hasta qué punto era limitada la concepción de la moralidad que brindó Bentham. Al suponer que la finalidad más elevada de la vida era el placer, aparentemente dejó de lado todo tipo de cosas que cualquiera de nosotros consideraría como inherentemente valiosas, tales como el conocimiento, el honor y los logros; Bentham propuso (éste es el cargo que le imputa Mill) «una doctrina digna sólo de cerdos». El propio Bentham, de un modo espléndidamente igualitario, se hizo

cargo de la acusación: « Dejando a un lado los prejuicios —declaró—, la petanca es un juego tan valioso como las artes y las ciencias de la música y la poesía» . Dicho de otro modo, si un juego popular produce una mayor cantidad general de placer, el juego es asimismo más valioso que las actividades más refinadas del intelecto.

A Mill le incomodaba la conclusión directa de Bentham y aspiraba a modificar el utilitarismo para evitar las críticas de sus adversarios. A las dos variables de Bentham para medir el placer —la duración y la intensidad—, Mill añadió una tercera —calidad— mediante la que introducía una jerarquía de placeres elevados y bajos. De acuerdo con esta distinción, algunos placeres, tales como los del intelecto y las artes, son más valiosos por naturaleza que los meramente físicos, y al atribuirles mayor peso en los cálculos de placer, Mill pudo concluir que « la vida del insatisfecho Sócrates es mejor que la de un tonto satisfecho» . Pero esta adaptación tuvo su precio. En última instancia, uno de los atractivos aparentes del esquema de Bentham —su simplicidad— había perdido fuerza, aunque la operación del cálculo de la felicidad implica una considerable dificultad en cualquier caso. Pero existe un problema más serio: la noción de Mill de distintos tipos de placer parece exigir algún criterio distinto del placer para poderlos distinguir. Si la idea de utilidad la constituye algo más que el placer, Mill debería poder hacer frente a problemas como el que plantea Nozick, aunque entonces la cuestión tal vez sea hasta qué punto su teoría sigue siendo estrictamente utilitarista.

La idea en síntesis: ¿basta con la felicidad?

«Sabes que Cristina quiere matar a tu amiga María, que acaba de levantarse de su silla en el bar. Cristina se acerca y te pregunta si sabes dónde está María. Si le dices la verdad, Cristina encontrará a María y la matará. Si mientes y le dices que viste a María irse hace cinco minutos, Cristina estará fuera de escena y María podrá huir. ¿Qué debes hacer, decir la verdad o mentir?»

Parece absurdo preguntarlo. Las consecuencias de decir la verdad son fatales. Naturalmente hay que mentir (una mentirijilla, piensas, por una buena causa). Pero para Immanuel Kant —uno de los filósofos más influyentes y, según algunos, el filósofo más importante de los últimos 300 años— ésta no es una respuesta correcta. No mentir es, para Kant, un principio fundamental de la moral, un «imperativo categórico»: algo que estamos obligados a hacer, incondicionalmente y a pesar de las consecuencias. La implacable insistencia en el deber, junto con la noción del imperativo categórico que subyace al mismo, constituye la piedra angular de la ética kantiana.

Imperativos hipotéticos versus imperativos categóricos Para explicar qué es un imperativo categórico, Kant empieza contándonos lo que no es, contraponiéndolo al imperativo *hipotético*. Supongamos que os digo qué hacer dando una orden (un imperativo): «¡No fuméis!». Implícitamente, existe una serie de condiciones que van asociadas a este mandato: «si no quieres echar a perder tu salud», por ejemplo, o «si no quieres derrochar el dinero». Naturalmente si no te preocupa la salud o el dinero, la orden carece de cualquier valor y no te sientes obligado a obedecer. Pero con un imperativo categórico no existen condiciones adicionales, ni implícitas ni explícitas. «No mientas» o «no mates» son mandatos que no dependen de ninguna finalidad o deseo que debieras o no debieras tener, de modo que hay que obedecer como una cuestión de deber, absoluta e incondicionalmente. Un imperativo categórico de esta naturaleza, a diferencia de lo que ocurre con un imperativo hipotético, constituye una ley moral.

Cronología

c. 30 d. C.	1739
La regla de oro	La guillotina de Hume

Desprecia las consecuencias

La ética kantiana, el sistema moral paradigmáticamente deontológico o fundamentado nos en el deber, ha tenido una profunda influencia en los sucesivos autores éticos, que han desarrollado con avidez sus ideas o las

han rechazado de forma enérgica. El meollo de la situación de Cristina y María fue planteado por Kant, e implacablemente asediado por su arsenal categórico: Kant insistió en que decir la verdad siempre es un deber moral, incluso cuando encontramos ante un asesino. En su infatigable insistencia en el deber por el deber, con independencia de cualquier consecuencia, previsible o imprevista, Kant perfila el modelo concebible más opuesto a los sistemas morales basados en las consecuencias.

Desde el punto de vista kantiano, tras cada acción subyace una regla de conducta, una máxima. Pero tales máximas pueden tener la forma de imperativos categóricos sin ser leyes morales, cuando no pasan una prueba, que es en sí misma una forma suprema o predominante del imperativo categórico:

> Actúa sólo de acuerdo con una máxima que puedas considerar simultáneamente como una ley universal.

Dicho de otro modo, una acción es moralmente aceptable sólo si se adecúa a una regla que sea aplicable de modo consistente y universal a ti mismo y a los otros (en efecto, se trata de una variación de la regla de oro; véase página 80). Por ejemplo, podríamos proponer la máxima de que mentir es aceptable. Pero sólo es posible mentir a partir de un fondo (de algún nivel) de discurso verdadero —si todo el mundo mintiera en todo momento, nadie creería a nadie— y precisamente por ello resultaría contraproducente y en buena medida irracional desear que la mentira se convirtiera en una regla universal. Asimismo, robar presupone un contexto de propiedad, pero la integridad del concepto de propiedad se hundiría si *todos* robáramos; faltar a una promesa presupone la institución aceptada por todos de mantener las promesas; etcétera.

1781 El imperativo categórico

1785 Medios y fines

El filósofo rumiante

Durante mucho tiempo Kant ha sido caricaturizado como el filósofo consumado y arquetípico, enclaustrado en su torre de marfil y sumido en sus meditaciones metafísicas teutonas. A esta imagen contribuye el

hecho de que Kant pasara toda su dilatada vida como un académico soltero en Königsberg: según parece, jamás puso un pie fuera de su ciudad natal.

La rigurosa austeridad de su filosofía y la endemoniada dificultad del lenguaje en que está escrita, intensifican los tonos oscuros del retrato. Pero alguna que otra vez Kant se desvía del camino, y proporciona munición a sus agresores; uno de los casos más célebres son sus meditaciones sobre el amor erótico, que (como el filósofo Simon Blackburn ha señalado) más bien parecen una descripción de una violación en grupo:

«En sí mismo constituye una profundas y oscuras degradación de la naturaleza humana; pues en cuanto una persona se convierte en el objeto del apetito de la otra, todos los motivos del vínculo moral dejan de funcionar, porque como objeto de apetito de otro una persona se convierte en una cosa y puede ser tratado y usado como tal por cualquiera».

Sea como fuere, por más razones que existan para la caricatura, el veredicto definitivo es que Kant es uno de los pensadores más originales e influyentes de la historia de la filosofía; un pensador cuya impronta indeleble puede hallarse tanto en la ética moderna, como en la epistemología o la metafísica.

El requisito de la universalidad descarta así un determinado tipo de conductas por razones lógicas, pero parecen existir muchas otras que podrían unlversalizarse, aunque no las consideráramos morales. No parece inconsistente ni irracional querer que reglas como «vela siempre por tus propios intereses», «falta a las promesas siempre que hacerlo no destruya la institución de la promesa» se convirtieran en leyes universales. ¿Cómo afronta Kant este peligro?

«**Dos cosas llenan el espíritu de una admiración y una veneración que no hace más que aumentar cuanto más frecuente y regularmente reflexionamos sobre ellas: el cielo estrellado sobre mi cabeza y la ley moral en mi interior.**»

La autonomía y la razón pura Los requisitos del imperativo categórico imponen una estructura racional a la ética kantiana, de modo que la labor consiste entonces en pasar del marco lógico a un contenido moral concreto (para explicar cómo puede «la razón pura», sin base empírica, dar forma y dirigir la voluntad del agente moral). La respuesta se encuentra en el valor inherente del comportamiento moral en sí mismo: valor basado en «el único principio supremo de la moralidad»,

Immanuel Kant, 1788 la libertad o la autonomía de la voluntad, que obedece a leyes que se impone ella misma. La suprema importancia asociada a los que actúan autónoma y libremente puede percibirse en la segunda gran formulación del imperativo categórico:

> Actúa de tal modo que nunca trates a la humanidad, en tu persona o en la persona de otro, como un simple medio sino siempre también como un fin.

Una vez que el inestimable valor de la propia capacidad para actuar moralmente se reconoce, es necesario ampliar su alcance a la acción de los otros. Tratar a los demás como un simple medio para propiciar los propios intereses degrada o destruye su capacidad de actuar, de modo que las máximas que son útiles para uno mismo o perjudiciales para los otros contravienen esta formulación del imperativo categórico y no tienen valor de leyes morales. En esencia, lo que encontramos aquí es el reconocimiento de que existen derechos básicos que pertenecen a las personas en virtud de su humanidad y que no pueden ser negados: he aquí la dimensión profundamente humana e ilustrada de la ética kantiana.

La idea en síntesis: el deber a cualquier precio

19 La regla de oro

«La cuestión es si a todos los americanos deben reconocérseles los mismos derechos y las mismas oportunidades, si vamos a tratar a nuestros conciudadanos americanos como nos gustaría que nos trataran a nosotros. Si un americano, por el hecho de tener la piel oscura, no puede tomar su almuerzo en un restaurante, ni llevar a sus hijos a la mejor escuela posible, ni votar a los funcionarios públicos que lo representarán, si, en resumidas cuentas, no puede disfrutar plena y libremente de la vida que todos nosotros queremos, ¿a quién de nosotros le gustaría entonces que le cambiaran el color de la piel y ocupar su lugar? ¿A quién de nosotros le complacería la paciencia y la espera?»

«**No hagas daño a nadie y nadie te hará daño.**»

Mahoma, *c*. 630

En junio de 1963, en una época en que la tensión y el odio raciales en Estados Unidos estaban desembocando en violencia y en constantes manifestaciones, el presidente John F. Kennedy habló al pueblo americano oponiéndose con firmeza a la segregación y a la discriminación raciales. Su discurso se centraba en la apelación a uno de los principios morales más fundamentales y ubicuos, la llamada « regla de oro» . Encapsulada en el dicho « No hagas a los otros lo que no te gustaría que te hicieran a ti» , la noción que subyace a esta regla parece ser fundamental para el sentido moral de la mayoría de los seres humanos y la encontramos con distintas formas en prácticamente todas las tradiciones morales y religiosas.

Muy pocos filósofos han conseguido evitar invocar la regla de oro o por lo menos señalar su relación con los principios de sus propias teorías. Aunque Kant afirmara que la regla de oro carecía del rigor necesario para establecerse como ley universal, existen evidentes similitudes con ella en la formulación más célebre de su imperativo categórico: « Actúa sólo de acuerdo con una máxima que puedas considerar simultáneamente como una ley universal» (véase la página 77). En el otro extremo del espectro filosófico, J. S. Mill reivindicaba la regla de oro para el utilitarismo, afirmando que « En la regla de oro de Jesús el Nazareno, encontramos el espíritu propio de la ética de la utilidad» (véase página 73). Podemos hallar un ejemplo más reciente en el prescriptivismo, la teoría ética desarrollada por R. M. Hare, quien propone que la noción de « lo universalizable» —evidentemente una variante de la regla de oro— constituye una propiedad esencial de los juicios morales.

Cronología

c. 30 d. C.	1739
La regla de oro	La guillotina de Hume

Oportunistas e hipócritas

Los oportunistas son primos hermanos de quienes se burlan de la regla de oro (todos aquellos a los que les gusta hacer cosas que no les gusta que les hagan): su propósito consiste en disfrutar del beneficio de lo que se ha hecho pero sin incurrir en los riesgos del hacer. Por ejemplo, los trabajadores que no se sindican pero se benefician de un aumento de salario gracias a las acciones del sindicato; o los países que no toman ninguna medida para reducir sus emisiones de carbono pero se benefician de las acciones colectivas internacionales para reducir el calentamiento global. En estos casos el problema consiste en que tal vez ser un oportunista sea racional individualmente, en atención sólo a nuestro propio interés, pero si muchas personas razonaran del mismo modo no se conseguiría ninguno de los beneficios deseados. ¿Es adecuado entonces usar la coerción? ¿Es adecuado respetar a los miembros del sindicato que abogan por la afiliación obligatoria, o promover acuerdos internacionales vinculantes, respaldados por la amenaza de sanciones o de otras medidas de fuerza?

Otros parientes cercanos de los que no respetan la regla de oro son los hipócritas, quienes se toman la libertad de no comportarse como deberían al no practicar lo que predican: el párroco adúltero que exalta la santidad del matrimonio; el político que acepta un soborno mientras abomina de la indecencia financiera. Como ocurre con todas las violaciones de la regla de oro, la principal objeción en estos casos es la incoherencia: entre las opiniones que sostiene la gente y las convicciones que parece sugerir su comportamiento; entre la importancia que pretenden atribuir a ciertas proposiciones y la indiferencia que cabe deducir de sus acciones.

1781 El imperativo categórico

1974 La máquina de la experiencia

«**Así que trata a los demás como te gustaría que te trataran a ti, pues ello resume la ley y los profetas.**»

Jesucristo, *c*. 30 d. C.

Observadores ideales y espectadores imparciales

La universalidad de la regla de oro —la razón de que se dé de una forma u otra en prácticamente todos los sistemas éticos filosóficos o religiosos— se debe en parte a su pura generalidad. Así, de acuerdo con los gustos y las necesidades particulares, se considera que sus rasgos dominantes son, entre otros, la reciprocidad, la imparcialidad y la universalidad. El carácter proteico de la regla ha implicado (casi) siempre la posibilidad de que aparezca de formas muy variadas en muchos sistemas distintos. Una de sus versiones más influyentes es la del «observador ideal». El supuesto en este caso consiste en que distintos factores como la ignorancia, la toma de partido por los amigos y la falta de simpatía hacia los extraños corrompen nuestros instintos puros e indomables. Como antídoto, se introduce a un observador ideal (o idealizado) cuya visión no está enturbiada por estas interferencias, y que así proporciona un criterio moral adecuado.

Uno de los desarrollos más célebres de esta noción es el «observador imparcial e informado» propuesto por el filósofo y economista escocés Adam Smith en su *Teoría de los sentimientos morales de* 1759. El espectador es la voz de la conciencia interior, «el hombre que hay en el interior de nuestros corazones, el gran juez y árbitro» de nuestra conducta; su jurisdicción se basa «en el deseo de poseer las cualidades y realizar los actos que amamos y admiramos en los demás; y en el horror de poseer las cualidades y realizar los actos que odiamos y despreciamos en los demás».

«La regla de oro es un buen criterio que puede mejorar haciendo por los otros cualquier cosa razonable que quisieran que hiciéramos por ellos.»

Karl Popper, 1945

Comprender la regla de oro A pesar de su atractivo inmediato, no está claro hasta qué punto la regla de oro puede brindarnos de hecho una guía práctica. Su perfecta simplicidad, aunque constituye uno de sus principales atractivos, la hace muy vulnerable a las críticas. Las personas disfrutan de modos muy distintos; la mayoría no masoquista debería desconfiar del masoquista que se atuviera rigurosamente a la regla. Además, cuando intentamos definir y precisar la regla, corremos el riesgo de sustraerle

«No desees a los otros lo que no deseas para ti mismo … Si deseas reconocimiento, ayuda a los otros a conseguirlo; si deseas éxito, ayuda a los otros a alcanzarlo.»

Confucio, *c*. 500 a. C.

toda su fuerza. Podríamos desear especificar el contexto y las circunstancias de aplicación de la regla, pero cuando especificamos tanto la regla empieza a perder la universalidad, que constituye una parte muy importante de su interés. La regla de oro encierra en su seno una demanda de consistencia, pero el egoísta puede perseguir de forma consistente su propio interés, y el hecho de que recomiende a los demás actuar del mismo modo no comporta ninguna inconsistencia.

En vez de considerar la regla de oro como una panacea moral (como algunos han pretendido), es más provechoso considerarla como un ingrediente esencial, una parte necesaria de los fundamentos de nuestro pensamiento ético: una demanda no sólo de consistencia sino de imparcialidad; la exigencia de intentar ponemos en el lugar del otro imaginariamente, de dispensar a los otros el mismo respeto y comprensión que esperamos recibir de ellos. Como tal, la regla de oro es un valioso antídoto a una especie de miopía moral que a menudo afecta a las personas cuando están en juego sus propios intereses.

La idea en síntesis: no hagas lo que no te gustaría que te hicieran

«El nivel del agua ya llega por encima del pecho de los espeleólogos y sube cada vez más deprisa. Si el equipo de rescate no interviene enseguida, los ocho hombres morirán en menos de media hora. Pero ¿qué puede hacer el equipo de rescate? No hay modo de sacar a los hombres a tiempo, ni de detener el flujo de agua. La única opción es desviar el agua hacia una gruta menor cercana. Pero en ella se encuentran los dos espeleólogos que se han separado del grupo principal atrapado: y están a salvo, esperando pacientemente que los rescaten. Desviar el curso del agua inundará la gruta menor en unos minutos y los dos hombres que están allí se ahogarán. ¿Qué debe hacer el equipo de rescate, pues? ¿Cruzarse de brazos y dejar morir a los ocho hombres, o salvar sus vidas al precio de las de sus dos colegas espeleólogos?»

Un dilema horrible y difícil de resolver. Supongamos que sólo existen dos posibilidades: desviar el flujo de agua, lo cual representa una intervención deliberada que causa la muerte de dos personas que de otro modo conservarían sus vidas; y sentarse y no hacer nada, lo cual permite que mueran ocho personas que podrían haberse salvado. Aunque la última opción es más grave desde el punto de vista del número de vidas perdidas, a muchos de nosotros nos da la impresión de que es peor actuar de un modo que provoca la muerte de alguien que dejar que mueran por efecto de nuestra pasividad. Como era de prever, la presunta diferencia moral entre lo que hacemos y lo que permitimos que ocurra —la llamada doctrina de los actos y las omisiones— divide a los autores que teorizan sobre cuestiones éticas. Quienes insisten en que el valor moral de un acto debería juzgarse estrictamente a partir de sus consecuencias rechazan la doctrina; en cambio, suelen apelar a los filósofos que hacen hincapié en la propiedad intrínseca de determinado tipo de acciones, y en nuestro deber de llevarlas a cabo con independencia de sus consecuencias (véase la página 76).

Cronología

c. 300 a. C.	*c.* 1260	1739
El problema del mal	Actos y omisiones La guerra justa	La guillotina de Hume

El principio del doble efecto

Al evaluar moralmente una acción, la intención del agente se considera crucial. Nuestras acciones pueden ser censurables incluso si las malas consecuencias que se derivan de ellas no son intencionadas (podrían evidenciar negligencia, por ejemplo), pero las mismas acciones tenderán

a ser valoradas con mayor severidad si las consecuencias son intencionadas. El principio del doble efecto (íntimamente vinculado con la doctrina de los actos y las omisiones) ataca la idea de separar las consecuencias deseadas de una acción de las consecuencias meramente previsibles. Una acción que tiene resultados tanto buenos como malos puede entonces justificarse desde un punto de vista moral si se llevó a cabo con la intención de conseguir los buenos resultados, mientras que los malos resultados eran previsibles pero no deseados. El argumento se ha puesto a prueba en casos como éstos:

• La vida de una madre se salva mediante una intervención quirúrgica en que se le extrae (y se mata) un feto: salvar la vida de la madre era el objetivo que se deseaba conseguir; matar el feto era un efecto previsible pero no deseado.

• Se suministran drogas mortíferas a pacientes terminales: la intención es paliar el dolor; el efecto no deseado pero conocido es que sus vidas se acortan.

• Se bombardea una fábrica de armamento del enemigo: la intención es destruir la fábrica; la consecuencia no deseada pero previsible (o «los daños colaterales») es que muchos civiles que viven en las inmediaciones mueren.

En todos estos casos la idea del doble efecto interviene para reforzar el supuesto de que las acciones en cuestión son moralmente defendibles. Recurren a menudo a la doctrina los pensadores favorables a una concepción de la moralidad absolutista o basada en el deber (deontológica) para explicar casos en que los deberes entran en conflicto y los derechos son aparentemente quebrantados. La consistencia o debilidad del principio depende de la distinción entre intención y previsión; y se ha discutido mucho acerca de si semejante distinción está en condiciones de soportar la responsabilidad que se le exige.

1785	1958	1971
Medios y fines	Más allá del sentido del deber	El principio de la diferencia

Jugar a ser Dios Por más firmes que sean nuestras convicciones en estos casos, a medida que examinamos con atención la distinción se va haciendo cada vez más imprecisa. La mayor parte de su atractivo, sobre todo en un asunto como la vida y la muerte, se debe a nuestro temor de «estar jugando a ser Dios» al hacer activamente las cosas: al decidir quién debe vivir y quién debe morir. Pero

¿en qué sentido propiamente moral « sentarse y no hacer nada» es, de hecho, no hacer nada? No actuar es una decisión, tanto como actuar, de modo que en tales casos no parece que tengamos otra elección que jugar a ser Dios. ¿Debemos mirar con peores ojos a los padres que escogen ahogar a sus hijos en el baño que a los que deciden no alimentarlos y dejarlos morir lentamente de hambre? Las sutiles distinciones entre matar y dejar morir parecen grotescas en estos casos, y nos costaría mucho decir que la « omisión» resulta en algún sentido menos reprobable que la « acción» .

Tomás de Aquino sobre la autodefensa

El planteamiento de lo que más tarde terminó conociéndose como la doctrina del doble efecto se atribuye generalmente al filósofo del siglo XIII Tomás de Aquino. Al discutir la justificación moral de matar en defensa propia, estableció distinciones que se encuentran muy próximas a las que recogen las modernas definiciones legales. El clásico planteamiento de la doctrina se encuentra en la *Summa Theologica* de Tomás de Aquino:

« Nada impide que un acto tenga dos efectos, y que sólo uno de ellos sea deseado, mientras que el otro no sea intencionado … el acto de la defensa propia puede tener dos efectos, uno es salvar la propia vida y el otro es dar muerte al agresor. Por eso, puesto que la intención es salvar la propia vida, este acto no es ilícito, ya que es natural que cualquiera quiera preservar su existencia tanto tiempo como sea posible. Y asimismo, aunque proceda de una buena intención, un acto puede devenir ilícito si es desproporcionado con relación al fin que persigue alcanzar. Porque si un hombre, en defensa propia, usa más violencia de la necesaria, también será un acto ilícito el suyo: mientras que si rechaza la fuerza con moderación su defensa será legítima» .

Enola Gay

¿Qué hubiera ocurrido si el bombardero B-29 *Enola Gay* no hubiera lanzado la primera bomba sobre Hiroshima el 6 de agosto de 1945? Todo parece indicar que aquella acción, a la invasión que siguió el lanzamiento de una segunda bomba sobre Nagasaki tres días después, precipitó el final de la segunda guerra mundial: Japón se rindió el día 14 de agosto.

Puede argumentarse que, a pesar del acto deliberado que causó una

cantidad atroz de muertes, se salvaron muchas más vidas puesto que se evitó la sangrienta japonesa. Así pues ¿estaba una segunda bomba sobre justificada la decisión de lanzar «la bomba»? Según el presidente Truman, «No hubo ninguna decisión que deba lamentarse».

La presunta distinción moral entre cosas hechas y cosas consentidas a menudo se invoca en cuestiones éticas médicas sensibles, como la eutanasia. En este caso suele establecerse una distinción entre la eutanasia activa, en la que los tratamientos médicos apresuran la muerte del paciente, y la eutanasia pasiva, en la que la muerte es consecuencia de la interrupción del tratamiento. Efectivamente, la mayoría de los sistemas legales (probablemente de acuerdo con nuestros instintos en este caso) recogen esta diferencia, pero sigue resultando muy difícil establecer cualquier distinción moral relevante entre, por poner un ejemplo, la administración de fármacos letales (un procedimiento deliberado) y la sustracción de fármacos para prolongar la vida (una omisión deliberada). La posición legal se basa en parte en la noción (fundamentalmente religiosa en origen) del carácter sagrado de la vida humana; pero al menos en lo que se refiere al debate sobre la eutanasia se trata sobre todo de una preocupación por la vida humana *per se*, con muy poca o ninguna consideración por su calidad o por las preferencias del ser humano cuya vida se discute. La ley tiene, pues, la extraña consecuencia de tratar con menos atención a un ser humano en un estado de extremo sufrimiento y angustia, que a un animal de compañía o de las granjas ganaderas en circunstancias similares.

La idea en síntesis: ¿hacer o no hacer?

21 Terrenos resbaladizos

El elevado terreno moral está rodeado de montañas, y las montañas implican laderas (muchísimas, y muy peligrosas también). En los debates públicos acerca de un buen número de asuntos sociales y políticos, el fantasma que aparece con mayor frecuencia y ahínco es el del terreno resbaladizo. La imagen es aparentemente tan sugestiva que se presenta a menudo sin demasiada comprobación y se acepta sin ofrecer demasiada resistencia. Además, aunque el deslizamiento hacia el terreno resbaladizo no es necesariamente ilícito, casi siempre se propone en los casos de asuntos con una considerable carga emotiva y, en muchos casos, el recurso es engañoso o elusivo.

La forma general del argumento del terreno resbaladizo está lejos de ser simple: si permites que se practique A (sea una práctica inocua o ligeramente discutible), ello conducirá inevitablemente a practicar Z (detestable e indeseable). Los terrenos resbaladizos pueden detectarse en un número apabullante de situaciones. He aquí algunos ejemplos clásicos:

• Permitir la eutanasia activa para que los enfermos terminales escojan el momento de su muerte nos abocaría inexorablemente a un ambiente de culpa en el que los ancianos aceptan «pasar en silencio», para hacer espacio, para reducir la carga de los costes y los cuidados a los jóvenes, etcétera.

• Permitid a los padres seleccionar el sexo de sus hijos, y dentro de nada querrán decidir cualquier otro atributo deseable y estaremos sumidos en la pesadilla de los niños de diseño.

• La legalización de drogas blandas como el cannabis fomentará la experimentación con drogas duras, y antes de damos cuenta las calles estarán forradas de las jeringuillas de los yonquis.

Cronología

c. 300 a. C.

La paradoja sorites

• La indulgencia con los delincuentes jóvenes los anima a cometer actos criminales más graves, y antes de darnos cuenta nuestras casas serán asediadas por jóvenes ladrones y asesinos.

Una característica común de todos estos argumentos es afirmar que entre A y Z se abre un terreno resbaladizo, y omitir los espacios intermedios que se abren entre B e Y. Por lo general, la omisión más notable es la más importante, una especie de justificación de la mencionada inevitabilidad de Z a partir de A. Se centra la atención en Z, que suele teñirse de los colores más siniestros, y con ello

se espera evitar cualquier discusión para que los méritos de practicar A pasen inadvertidos. En efecto, el argumento es sustituido por la retórica. La conveniencia, por ejemplo, de permitir a los padres escoger el sexo de sus hijos debería considerarse a partir de su propio valor, y si se considera objetable deberían existir buenas razones para impedirlo. Si se establece que la práctica es por sí misma inofensiva, podría ser importante considerar la presunta inevitabilidad de otras prácticas asociadas que sí son objetables. Pero sería difícil que se diera el caso, puesto que en la vida real, donde el peligro de los terrenos resbaladizos es genuino, normalmente siempre es posible plantear normas y pautas para prevenir patinazos no deseados.

El dominó, los icebergs y las fronteras Los terrenos resbaladizos no son los únicos accidentes de los que nos alertan los moralizadores públicos. El primer paso en falso en el terreno resbaladizo suele precipitarnos cuesta abajo hasta hundirnos en el bosque de otros peligros verbales, donde se oye el repiqueteo de las piezas de dominó que caen, y el ruido de bolas de nieve que cobran dimensiones monstruosas en su caída, y de esclusas completamente abiertas, y donde cada iceberg anuncia profundidades ocultas.

Hervir ranas

Los peligros de la cautela frente a los cambios políticos y sociales se ilustran a veces con la historia de la rana hervida: si la echamos al agua hirviendo no servirá de nada (de ahí la historia), pues inmediatamente saldrá de un salto; este inconveniente puede evitarse, sin embargo, echando la rana en agua fría y llevándola a ebullición lentamente. Del mismo modo, los liberales de alma anfibia argumentan que la erosión gradual de nuestras libertades civiles nos aboca a una pérdida acumulada (o a una «usurpación del poder») que habría podido evitarse firmemente con una sola intervención. La teoría sociopolítica es más plausible que la teoría de la rana; la falsedad de la última debería asumirse, no probarse.

Del mismo modo que al caer una pieza de dominó puede golpear sobre la siguiente e iniciar así una cadena de sucesivas caídas, con el **efecto dominó** se sugiere que el advenimiento de un acontecimiento particular no deseado desencadena una secuencia de acontecimientos similares asociados. El ejemplo

más conspicuo de esta idea es del año 1954, cuando surgió la «teoría del dominó» propuesta por el presidente de Estados Unidos Dwight Eisenhower para justificar la intervención estadounidense en Vietnam. De acuerdo con la teoría, permitir que un país cayera en manos de los comunistas llevaría inevitablemente a que ocurriera lo mismo en otros países asiáticos. En ese caso la primera pieza del dominó, Vietnam, había caído, pero si se conseguía que Camboya no lo hiciera la previsible propagación del comunismo por toda la región se evitaría; en este caso fue muy evidente que la supuesta inevitabilidad no era tal.

Una pequeña grieta en una piedra o en la madera puede ir abriéndose progresivamente si se introduce una cuña; del mismo modo, cuando se apela a la metáfora de la punta del iceberg se pretende sugerir que un cambio aparentemente insignificante, por ejemplo una norma o una ley, puede ser el principio de, o la excusa para, una reforma general. El supuesto de que podría prescindirse del derecho a juicio con jurado en los complejos casos de fraude es percibido por algunas personas como la punta del iceberg, pues sospechan que propiciará un progresivo recorte de derechos en otros ámbitos (tal vez en todos). Pero la sospecha de los teóricos sigue siendo una mera suposición hasta que las evidencias muestren que los políticos tienden a confirmarla.

El hocico del camello

Una variante de la metáfora del terreno resbaladizo muy divertida y exótica, supuestamente basada en una fábula oriental, nos permite vislumbrar de un modo encantador los peculiares peligros de la vida en una tienda (o en la piel de un camello). Las nefastas consecuencias de «dejar que el camello meta el hocico en la tienda» —especialmente porque la nariz es sin duda alguna la parte más desagradable del camello— las captó de un modo delicioso la poeta norteamericana del siglo XIX Lydia Howard Sigourney:

> Cuando a su tienda llegó un hombre por el trabajo estragado,
> con los brazos molidos, y el pensamiento embotado,
> a través del espacio de la ventana abierta
> ¡Qué vio! Un camello asomaba la testa.
> «Tengo la nariz helada», lloró lastimeramente,
> «ay, deja que un poco me la caliente».
> Como al meter el hocico nadie le dijo que no
> el largo y sinuoso cuello tras la cabeza siguió

igual que a la misa la Eucaristía
y cuando luego intensa lluvia caía
de un brinco entero todo su cuerpo metió.
Aterrado, su alrededor el hombre escrutó
y a su insolente invasor escudriñó
pues cuanto más cerca lo veía,
para aquel invitado habitación no había,
mas pasmado le oyó decir
« Si estás incómodo te puedes ir
pues de aquí tú ya no me mueves» .
Oh, qué frívolo corazón joven eres,
del pobre árabe mofarte no debes
pues la mala costumbre es, como verás,
la primera treta que sufrirás.
No la escuches, ni la mires, ni le sonrías jamás,
la oscura fuente antes de que brote ahogar deberás
y ni siquiera el hocico al camello consentirás.

El problema de dónde **establecer la frontera** habitualmente procede del afán de conocimiento acerca de cosas sobre las que no es posible ningún conocimiento, es decir, de la expectativa de un grado de precisión inadecuada en el contexto. Todos nosotros podemos estar de acuerdo, por ejemplo, en que sería un error permitir que millones de inmigrantes entraran en nuestro país cada año y en que, no obstante, es correcto permitir la llegada de algunos. ¿Dónde establecer la frontera? El hecho de que exista necesariamente un grado de vaguedad acerca de una decisión, o del contexto en el que se plantea, no significa que no pueda o no deba tomarse dicha decisión. Ése es exactamente el tipo de problema que ha dominado el debate sobre el aborto: muchos coinciden en que un embrión recién gestado y un bebé completamente desarrollado son distintos, aunque resulte complicadísimo (por imposible) establecer con exactitud el momento preciso en el que se produce ese cambio. Ello ocurre porque el desarrollo del feto es un proceso gradual, y cualquier punto en el que decidamos establecer la frontera será en alguna medida arbitrario. Lo cual no significa que cualquier punto sea tan

bueno (o malo) como otro, que no pueda establecerse ninguna frontera, o que cualquier frontera que establezcamos carezca de autoridad o de consistencia.

La idea en síntesis: si les das una mano se tomarán…

22 Más allá del sentido del deber

El 31 de julio de 1916, durante la batalla del Somme en el norte de Francia, el joven de 26 años James Miller, un soldado raso del Regimiento Real de Lancaster, «recibió la orden de llevar un mensaje importante atravesando el fuego cruzado para volver con la respuesta costara lo que costara. Lo obligaron a partir a campo abierto y al salir de la trinchera recibió inmediatamente un disparo en la espalda que lo atravesó y salió por el abdomen. A pesar de ello, con un coraje y un espíritu de sacrificio heroicos, oprimió con su mano la herida del estómago, entregó el mensaje, regresó tambaleándose con la respuesta, y se desplomó a los pies del oficial a quien se la entregó. Sacrificó su vida con una devoción y un sentido del deber extremos».

¿Cómo interpretar este tipo de comportamientos? Es evidente que las autoridades militares británicas durante la primera guerra mundial consideraron las acciones del soldado Miller excepcionales —incluso en una época en que se producían muchos hechos extraordinarios cotidianamente—, pues lo condecoraron con la Cruz de la Victoria «al valor más sobresaliente» (el texto citado procede del discurso del oficial de Miller en ocasión de la condecoración). Si Miller se hubiera arrastrado para volver a la trinchera inmediatamente después de recibir el disparo que podía acabar con su vida, difícilmente lo *culparíamos*, o diríamos que actuó de un modo *equivocado*, o que su acción fue *inmoral*. Igual que sus oficiales, tenderemos a considerar que los actos de Miller iban «más allá del sentido del deber» y que merecían una alabanza especial. En suma, seguramente elogiaríamos lo que hizo pero no le hubiéramos culpado si hubiera actuado de otro modo.

Cronología

c. 30 d. C.	*c.* 1260	1739
La regla de oro	Actos y omisiones	La guillotina de Hume

Actos heroicos En principio, nuestras intuiciones comunes armonizan con este tipo de afirmaciones. Parece normal concebir dos niveles distintos en la moralidad. En uno de ellos, se encuentran las cosas que todos nosotros estamos obligados moralmente a hacer: obligaciones básicas que son cuestión de deber y responden al patrón mínimo de moralidad común. A menudo se las considera negativamente, como obligaciones con las que es incorrecto no cumplir: no mentir, estafar, matar, etc. Suponemos que debemos cumplirlas y esperamos que los otros hagan lo propio.

Además de estos deberes morales comunes existen, en un nivel más elevado, los ideales morales. Suelen expresarse positivamente y no son normas cerradas: así,

mientras que existe el deber moral corriente de no robar a los otros, la gran generosidad con los demás es un ideal que en principio es ilimitado. Semejantes actos pueden ir más allá de lo que requiere la moral común y se inscriben en la categoría de los llamados « actos heroicos» : actos que resulta encomiable hacer pero que no es censurable omitir. Los actos heroicos son el territorio « de los héroes y de los santos» . Tales individuos consideran como un deber esos actos y se culpan a sí mismos si no obedecen a ellos, pero se trata esencialmente de un *sentido* personal del deber y no cabe juzgar a los demás a partir de ellos.

¿Actos heroicos?

« Imaginemos un pelotón de soldados que practican lanzamiento de granadas de mano; a uno de los soldados le resbala de las manos una granada que rueda hasta la zona donde se encuentra el pelotón; y otro sacrifica su vida echándose sobre la granada y protegiendo a sus colegas con su propio cuerpo … Si el soldado no se hubiera echado sobre la granada ¿hubiera sido culpable de no cumplir con un deber? Aunque de algún modo sea, sin duda, superior a sus compañeros, ¿podemos decir que ellos faltaron a su deber al no intentar sacrificarse? Y si ese soldado no lo hubiera hecho, ¿acaso alguien hubiera podido decirle “Deberías haberte echado sobre la granada”?»

1781	1785	1958	1974
El imperativo categórico	Medios y fines	Más allá del sentido del deber	La máquina de la experiencia

Esta historia se encuentra en « Santos y héroes» , un artículo muy importante que escribió en el año 1958 el filósofo inglés J. O. Urmson, y que ha orientado la mayoría de los debates filosóficos recientes en torno a los actos heroicos. Urmson identifica tres condiciones que debe reunir un acto para ser heroico: no puede tener por objeto un deber (común); debe ser encomiable y su omisión no puede implicar ninguna culpa. Urmson sostiene que el caso mencionado reúne los tres requisitos, razón por la cual se trata de un acto heroico.

La integridad moral

La idea de los actos heroicos pone de relieve la dimensión personal de la moral. Los héroes y los santos tienen un personal sentido del deber, de lo que les *corresponde* hacer, y por eso renuncian a su derecho a las

exenciones (como el peligro o la dificultad) que la mayoría de nosotros usamos para excusar no actuar en ciertos casos. La mayoría de las formas del utilitarismo son rigurosamente impersonales: en sus valoraciones morales consideran que toda vida (incluida la del que actúa) tiene el mismo valor, y tienden a subestimar la importancia de los propósitos y de los compromisos personales. El criterio utilitarista tiende a no tenerlos en cuenta cuando se trata de tomar decisiones morales, y en este sentido hay quienes consideran el utilitarismo como considerablemente deficiente en lo que toca a la explicación de las prioridades personales y del sentido de la integridad moral del sujeto de la acción.

¿Son opcionales los actos bondadosos? La categoría de acciones morales extraordinarias, no obligatorias, es interesante desde el punto de vista filosófico precisamente a causa de las dificultades que algunos sistemas éticos presentan para integrarlas. Tales sistemas se caracterizan por ofrecer una concepción de lo que es bueno a partir de la cual se define lo correcto y lo incorrecto por remisión a tal criterio. En consecuencia, la idea de que algo es reconocido como bueno y, aun así, no es exigible resulta muy complicada de explicar.

De acuerdo con el utilitarismo, al menos en sus versiones más comunes (véase la página 73), una acción es buena si incrementa la utilidad general (por ejemplo, la felicidad), y la mejor acción en cada situación es la que resulta de mayor utilidad. Dedicar una parte importante de nuestro dinero a donativos para los países en vías de desarrollo no sería considerado como una obligación moral; los otros pueden encomiarnos el hacerlo, pero no tienen por qué sentirse mal por no hacerlo. Dicho de otro modo, la caridad en este nivel tan generoso es extraordinaria. No obstante, desde el punto de vista utilitario, si semejante acción promueve la utilidad general (y evidentemente éste parece ser el caso), ¿cómo es posible que no se trate de algo exigible? También para la ética kantiana este tipo de actos heroicos resulta problemático. Kant establece el valor moral más elevado en la capacidad de actuar del sujeto (véase la página 76). Si ello es así, ¿cómo es posible que exista algún límite para lo que mejora o estimula esa capacidad de actuar?

Este tipo de conflictos entre las teorías éticas y nuestro sentido moral común resultan muy problemáticos, especialmente para las teorías éticas. Los utilitaristas radicales defenderían (es lo que hacen algunos) que debemos aceptar todas las implicaciones de sus teorías —negar, en efecto, que existan acciones heroicas— y cambiar nuestro modo de vida para adecuarla a ellas. Pero semejantes propuestas de reforma extrema, que se oponen a los principios de la

moralidad común, y califican la mayoría de nuestros actos como moralmente insuficientes, tienen más probabilidades de disuadir a la gente que de cautivarla. Por lo general, la mayoría de los teóricos intentan justificar o minimizar los conflictos aparentes. Una estrategia habitual consiste en apelar a alguna forma de excepción o de excusa (por ejemplo, una dificultad excesiva o un peligro) que permita a la persona actuar de determinado modo que, de lo contrario, resultaría obligatorio. Pero aunque esta estrategia brinde una determinada salida del atolladero, tiene su coste. Porque la consideración de algunos factores personales defrauda la universalidad que suele considerarse indispensable en el terreno moral (véase la página 80). Otra alternativa consiste en proponer ideas como la doctrina del doble efecto y la distinción entre actos y omisiones (véase la página 85) para explicar cómo puede ser correcto atenerse a un patrón cuando existe otro disponible y aparentemente preferible. Pero ninguna de estas ideas elude la dificultad y, en cualquier caso, muchos de nosotros podemos sentir que la verosimilitud de una teoría disminuye cuando está repleta de notas a pie de página y de otras reservas.

La idea en síntesis: ¿todos deberíamos ser héroes?

23 ¿Es malo tener mala suerte?

«Dos amigos, Bell y Haig, pasan la tarde juntos en el bar. Cuando llega la hora de cerrar, con una o dos cervezas más de la cuenta en el cuerpo, los dos se dirigen tambaleándose hacia sus coches para ir a casa. Bell llega a casa sin problemas, como muchas otras veces, se desploma en la cama, y se levanta al día siguiente con una ligera resaca. Haig —tan curtido y experto en la conducción con unas cuantas copas de más como Bell— regresa a casa relajadamente hasta que, de pronto, se le cruza en el camino un joven que atraviesa la calle frente a él. No hay tiempo de frenar, y el joven muere al instante. A Haig lo meten en la celda de una comisaría y al día siguiente se levanta con una ligera resaca y la certeza de que va a pasar años en la cárcel.»

¿Cómo interpretar el comportamiento de Bell y Haig? Naturalmente, la ley establece que el comportamiento de Haig es mucho más culpable: si a Bell lo hubieran pillado le habría caído una multa y le habrían retirado el carné durante unos meses: en cambio a Haig le caerá, con toda seguridad, una sentencia de prisión bastante larga. En este caso, el punto de vista legal reflejaría adecuadamente nuestro sentido moral. Consideramos que quien provoca la muerte a alguien a causa de un acto irresponsable es más culpable que otro que conduce traspasando, un poco, los límites legales de alcohol. No obstante, la única diferencia entre los dos conductores en este caso —el joven en medio de la calle — es la suerte. Los dos conductores actuaron de forma irresponsable, y uno de ellos tuvo mala suerte. De modo que la suerte es el único factor que parece explicar la sanción legal y la certeza moral de que el comportamiento de Haig es malo: pero la suerte es algo sobre lo que, por definición, no tenemos ningún control al actuar.

Cronología

c. 350 a. C.	*c.* 1260
La ética de la virtud	Actos y omisiones

La suerte moral La diferencia entre los dos casos parece contradecir una intuición muy común: la de que sólo es adecuado juzgar las cosas moralmente en la medida en que dependen de nosotros. Me parecería reprobable que me echaras el café encima de forma deliberada, pero seguramente no te culparía si el accidente tuviera lugar en un tren y lo hubiera provocado un súbito frenazo del maquinista. Puede plantearse lo mismo diciendo

«Si cometemos la negligencia de dejar el agua de la bañera correr con nuestro hijito dentro, advertiremos —mientras brincamos escaleras arriba en dirección al baño— que si nuestro hijo se ha ahogado habremos hecho

que no debería juzgarse a dos personas de un modo distinto a menos que las diferencias se deban a factores que dependen de ellos. Si al golpear la pelota un golfista la lanza contra el público, y golpea a un espectador y lo mata, no lo culparemos más, en caso de que lo culpemos de algo, que a otro jugador que hace un lanzamiento similar pero sin consecuencias funestas (aunque muy distinto es cómo se siente *consigo mismo* el desdichado jugador).

algo espantoso, mientras que si no ha ocurrido, simplemente se tratará de un descuido.»

Thomas Nagel, 1979

Pero si trasladamos este modo de pensar al caso de Bell y Haig, parece que deberíamos juzgarlos del mismo modo. Entonces ¿deberíamos juzgar más severamente a Bell por el perjuicio que su comportamiento irresponsable habría podido causar? ¿O deberíamos ser más indulgentes al juzgar a Haig, puesto que no estaba haciendo nada peor que muchas otras personas y, simplemente, tuvo mala suerte? Es evidente que podemos seguir ateniéndonos a nuestro juicio inicial: los dos casos deben considerarse distintos en atención a sus distintas consecuencias. Pero si mantenemos nuestra posición tendremos que modificar nuestra idea sobre la importancia del control: estaremos obligados a concluir que la moralidad no es inmune a la suerte: paradójicamente debería existir algo como la «suerte moral». Se diría que, después de todo, la suerte puede convertirnos en malvados.

1739	1789	1976
La guillotina de Hume	Teoría del castigo	¿Es malo tener mala suerte?

¿O es mala suerte ser malo? La cuestión de si existe la suerte moral (es decir, de si los juicios morales están determinados, al menos en parte, por factores aleatorios ajenos a nuestro control) ha sido el objeto de buena parte de las discusiones filosóficas recientes. El debate se recrudece cuando hace acto de presencia la llamada «suerte resultante», en casos como el de Bell y Haig, donde el azar resultante de una acción parece afectar a nuestra evaluación de la misma. Pero existen otros tipos de suerte que pueden intervenir y, de hecho, el problema puede llevarnos mucho más lejos.

A partir de casos parecidos al de Bell y Haig resulta tentador responder que las *intenciones* de quien actúa —no las consecuencias de esas intenciones— son lo que debería considerarse cuando prodigamos alabanzas o cuando dictaminamos culpabilidad. Bell y Haig tienen la misma intención (ninguno de los dos pretende matar a nadie) y, por lo tanto, deberían (poder) ser juzgados del mismo modo. Pero ¿hasta qué punto tenemos control sobre nuestras intenciones? Tenemos las intenciones que tenemos a causa del tipo de persona que somos, pero existen innumerables factores (que caen bajo la categoría general de «suerte

constitutiva») que nos configuran como personas y sobre las que no tenemos control. Nuestro carácter es el producto de una combinación sumamente compleja de factores ambientales y genéticos sobre los que tenemos muy poco o ningún control. ¿Hasta qué punto deberíamos ser juzgados por acciones e intenciones que surgen de forma natural de nuestro carácter? Si no puedo evitar ser cobarde o egoísta —si se da el caso de que está «en mi naturaleza» serlo—, ¿es acaso justo culparme o criticarme por huir del peligro o por pensar demasiado en mis propios intereses?

El lugar y el momento equivocados

Sólo podemos desplegar los aspectos buenos y malos de nuestros caracteres cuando las circunstancias nos proporcionan la ocasión de hacerlo: todos nosotros estamos a merced de la «suerte circunstancial». No es posible exhibir tu gran generosidad natural si careces de los recursos para ser generoso, o los beneficiarios potenciales con quien serlo. Podríamos sospechar que nunca habríamos caído en la depravación a la que llegaron los soldados nazis de Auschwitz, pero naturalmente nunca podremos estar del todo seguros. Todo lo que podemos afirmar con certeza es que somos muy afortunados de no habernos encontrado nunca en semejante situación. Así pues, ¿el soldado nazi tuvo la mala suerte de haber tenido que encontrarse en la situación? ¿Tuvo la mala suerte de ser malo?

«**El escepticismo en torno a la autonomía de la moralidad con respecto a la suerte no puede dejar el concepto de moralidad intacto … [El escepticismo] nos dejará con un concepto de la moralidad, pero será más irrelevante de lo que solemos considerar; y no será nuestro, pues una de las cosas particularmente importantes de nuestro concepto de moralidad es la importancia que le atribuimos.**»

Bernard Williams, 1981

Es posible desplazar los límites de la suerte cada vez más atrás. Si consideramos otro tipo de suerte —la circunstancial— advertimos hasta qué punto una evaluación sobre la maldad moral puede depender del estar en el lugar y el momento equivocados. Si extraemos la conclusión lógica, el debate acerca de si existe algo como la suerte moral linda con el problema de la voluntad, y plantea

los mismos problemas: en última instancia, ¿hay algo de lo que hacemos que hagamos libremente? ¿Y si no existe libertad, puede existir responsabilidad? ¿Y sin responsabilidad, qué justificación existe para la culpa y el castigo (véase la página 196)?

Las intuiciones comunes sobre la suerte moral están lejos de ser homogéneas o consistentes. Esta incertidumbre se refleja en cierta polarización de las posiciones filosóficas en torno al asunto. Algunos filósofos niegan que exista algo como la suerte moral, y tratan de explicar o de encontrar una justificación cuando aparece de forma manifiesta en nuestro discurso moral común. Otros aceptan la existencia de la suerte moral y prosiguen a partir de ahí evaluando si (y en qué medida exacta) ello nos obliga a reformar o revisar el modo de juzgar moralmente. Lo que está en juego es mucho: pueden verse afectados algunos supuestos básicos acerca del modo en que conducimos moralmente nuestras vidas; y una vez más no existe el menor signo de consenso.

La idea en síntesis: ¿la fortuna favorece el bien?

Durante más de 400 años, los filósofos morales se han fijado principalmente en las acciones, no en los actores, en el tipo de cosas que deberíamos hacer en vez de en el tipo de persona que deberíamos ser. La principal tarea del filósofo ha consistido en descubrir y explicar los principios en los que se basa el deber moral y en formular normas que orienten nuestro comportamiento de acuerdo con tales principios.

Son muchas las propuestas sobre la naturaleza de los principios mismos, desde la ética basada en el deber de Kant hasta el utilitarismo consecuencialista de Bentham y Mill. Sin embargo, todas ellas comparten el supuesto de que la justificación de las acciones es más fundamental que el carácter de los actores, que se considera secundario o meramente instrumental. Pero la virtud no ha sido siempre una sirvienta del deber o de algún otro bien superior.

Hasta el Renacimiento, con los primeros atisbos de la revolución científica, la influencia más importante de la filosofía y la ciencia procedía de los grandes pensadores de la Grecia clásica: Platón y, sobre todo, su pupilo Aristóteles. La principal preocupación de estos autores era la naturaleza y el cultivo del carácter bondadoso; la pregunta central no era: « ¿Qué es correcto hacer (en esta o en aquella circunstancia)?» , sino: « ¿Cuál es el mejor modo de vida?» . Dada la diferencia de prioridades, la naturaleza de la virtud, o la excelencia moral, era un asunto del mayor interés. La filosofía de Aristóteles quedó eclipsada durante algunos siglos, desde la época de Galileo y Newton, pues la atención se desplazó a las reglas y principios de la conducta moral. Sin embargo, a partir de mediados del siglo XX algunos pensadores empezaron a expresar su insatisfacción con respecto a la tendencia hegemónica en filosofía moral, y a interesarse de nuevo por el estudio del carácter y de las virtudes. Este movimiento reciente en la teoría moral, inspirado sobre todo en la ética aristotélica, se conoce con la etiqueta de « ética de la virtud» .

Cronología

c. 440 a. C.	c. 350 a. C.
La carne de un hombre...	La ética de la virtud

La ética y la moral

A primera vista existe tal abismo entre el modo en que Aristóteles concebía la ética y la concepción adoptada por la mayoría de los filósofos actuales, que se ha llegado a sugerir la necesidad de adaptar la

terminología para dar cuenta de tal diferencia. Se ha propuesto que el término «moralidad» debería restringirse a sistemas como el de Kant, en el que el interés se centra en los principios del deber y las normas de conducta: mientras que el término «ética» (derivado de la palabra griega para referirse a «carácter») debería reservarse a las concepciones aristotélicas, que priorizan las disposiciones del que actúa y la sabiduría práctica (no sólo moral). Ha habido disputas sobre la utilidad de la distinción, que se considera a veces como una falsa oposición, por exageradamente radical, entre Aristóteles y los filósofos que se le oponen.

Los griegos y la virtud Según Aristóteles y algunos otros pensadores griegos, ser una buena persona, y diferenciar lo correcto de lo incorrecto, no es principalmente un problema de entender y aplicar determinadas reglas y principios morales. Más bien es cuestión de ser y convertirse en el tipo de persona que, gracias a la búsqueda de la sabiduría a través de prácticas y ejercicios adecuados, consigue comportarse por lo general de un modo correcto en las circunstancias adecuadas. En resumen, tener el tipo de disposiciones (naturales y adquiridas) y de carácter correctos para el tipo de comportamiento correcto.

Las disposiciones mencionadas son virtudes. Expresan o manifiestan la *eudaimonia*, que los griegos consideraban el mayor bien del hombre, y el objetivo último de la actividad humana. Traducido habitualmente por «felicidad», la *eudaimonia* es, de hecho, algo más amplio y dinámico, que capta mejor la idea de «colmar» o «llevar una vida buena (venturosa, exitosa)».

1739	1974	1976
La guillotina de Hume	La máquina de la experiencia	¿Es malo tener mala suerte?

Los griegos solían hablar de cuatro virtudes fundamentales: el coraje, la justicia, la templanza (autodominio) y la inteligencia (la sabiduría práctica). Pero tanto para Platón como para Aristóteles la doctrina crucial es la llamada «unidad de las virtudes». Dada la constatación de que, en parte, una persona buena es la que sabe cómo responder sensiblemente a la demanda de diferentes virtudes que a veces entran en conflicto, estos autores concluían que las virtudes son como las distintas caras de una sola piedra preciosa, de modo que no es posible poseer una virtud sin poseerlas todas. En Aristóteles, la posesión y el cultivo de las distintas virtudes implica que el hombre bueno posee una virtud que las engloba a todas, la «magnanimidad» (del latín, «dotado de una gran alma»). El *megalopsychos*

(« el hombre de gran alma») aristotélico es el arquetipo de la bondad y la virtud: el hombre de posición distinguida en la vida, y digno de grandes cosas; afanoso de ofrecer ayuda pero reticente a recibirla; dotado de un orgullo cabal y de una humildad moderada.

El justo medio

El justo medio es una noción central de la concepción aristotélica de la virtud. La doctrina es considerada a veces erróneamente como una apelación a la moderación, en el sentido de atenerse siempre a un modelo intermedio, pero no es ése su significado. Tal como queda claro en la cita, el significado se define estrictamente en relación con la razón. Por ejemplo: la virtud que entraña el justo medio entre la cobardía y la audacia es el coraje. Tener coraje no es sólo evitar los actos cobardes, como salir corriendo ante el enemigo; también es necesario evitar la imprudencia, la bravuconada temeraria, como empecinarse ataque fútil que nos perjudique a nosotros y a nuestros compañeros. El coraje depende del gobierno de nuestra razón sobre nuestros instintos más bajos e irracionales: lo crucial es que la acción sea apropiada a las circunstancias, es decir, que esté determinada por una sabiduría práctica capaz de responder sensiblemente a los hechos particulares de la situación.

« La virtud es entonces una disposición del carácter que tiene que ver con la elección, y que descansa en el medio definido por la razón. Es el justo medio entre dos vicios, uno por exceso y otro por defecto; y una vez más, es un medio porque los vicios, o bien nos hacen excedernos o bien quedarnos cortos respectivamente con respecto a lo que es correcto para las pasiones y para las en un acciones, mientras que la virtud sabe encontrar y escoger lo que se encuentra en el medio.»

Aristóteles, *c*. 350 a. C.

«La Bondad del hombre es el ejercicio activo de las facultades de su alma de acuerdo con la excelencia o la virtud … Además, esta actividad debe ocupar toda la vida: pues igual que una flor no hace la primavera, tampoco basta un día perfecto.»

Aristóteles, *c*. 350 a. C.

La jerarquía que implica la unidad de las virtudes lleva a Platón a la firme conclusión de que las diferentes virtudes son, de hecho, una y la misma, y de que se encuentran subsumidas bajo una sola virtud: el conocimiento. La idea de que la virtud es idéntica al conocimiento también le lleva a negar la posibilidad de la *akrasia*, la voluntad débil: para él era imposible «saber lo que es mejor y no obstante hacer lo peor»; los actos intemperados, por ejemplo, no eran consecuencia de la debilidad sino de la ignorancia. También Aristóteles, siempre afanoso de evitar contradecir las creencias tradicionales (*endoxa*), se opuso a la idea de que podemos actuar mal a sabiendas (a pesar de lo que muestra la experiencia).

Tanto para Platón como para Aristóteles, comportarse de un modo virtuoso era inseparable del ejercicio de la razón, de la elección racional; y Aristóteles elaboró esta idea en la doctrina del justo medio (véase el cuadro de la página anterior).

La idea en síntesis: lo que eres, no lo que haces

25 ¿Sufren los animales?

«—¡Ay, mi pierna! —gritó—, ¡ay, mi pobre espinilla! —y se sentó en la nieve para curarse con las dos patas delanteras.
—¡Pobre viejo Topo! —dijo amablemente la Rata—, me temo que hoy no es tu día, ¿verdad? Vamos a ver esa pata —prosiguió, agachándose para mirar de cerca.
—Sí, te has partido la espinilla, está claro. Espera que saque el pañuelo para vendarte.

—Debo haber tropezado contra alguna rama enterrada bajo la nieve o algún pedrusco —dijo penosamente el Topo—. ¡Ay de mí, ay!

—Es un corte muy limpio —dijo la Rata, examinando de nuevo la herida con atención—. Una rama o un pedrusco no hace una herida así…

—Qué más da lo que habría sido —dijo el Topo, olvidando la gramática a causa del dolor—. Duele igual, sea lo que sea.»

¿Sufren realmente los animales, o sólo sufren los ficticios, como el topo de *El viento en los sauces*? Podemos estar razonablemente seguros de que los animales no hablan, pero sabemos poco más que eso.

El modo en que respondemos al problema del sufrimiento de los animales, y al más amplio problema de su conciencia, tiene implicaciones directas en otras cuestiones candentes:

• ¿Es correcto usar millones de ratas, ratones, e incluso primates, en investigaciones médicas, pruebas de fármacos, etc.?

• ¿Es correcto envenenar, gasear y exterminar topos y otros animales considerados « plagas» ?

• ¿Es correcto sacrificar en mataderos a millones de animales como vacas o gallinas para proporcionamos alimento?

Cronología

c. 250 a. C.	1637	1739
¿Sufren los animales?	El problema del cuerpo y la mente	La guillotina de Hume

El giro lingüístico

Por analogía con nuestras mentes, podemos inferir muchas similitudes entre las experiencias conscientes de los humanos y las de (algunos)

animales, ¿pero hasta dónde podemos llegar a partir de ahí? La experiencia subjetiva de un animal debe estar íntimamente unida a su forma de vida y al entorno particular al que está adaptado evolutivamente; y como Thomas Nagel señaló, no tenemos la más remota idea de cómo sería ser un murciélago, o cualquier otro animal (véase la página 36). Este problema se hizo aún más los evidente a partir del « giro lingüístico» que terminó por dominar la mayor parte de la filosofía de la mente en el siglo XX. De acuerdo con este giro, nuestra vida mental está apuntalada o mediada por el lenguaje, y nuestros pensamientos son necesariamente representados en nuestro interior en términos lingüísticos, Tal concepción, aplicada sin más a los animales, nos obligaría a negar que puedan abrigar algún pensamiento. Las posturas se han ido moderando y, en la actualidad, la mayoría de los filósofos estarían dispuestos a aceptar que el resto de animales también tienen pensamientos, por más simples que sean.

La mayoría de los filósofos están de acuerdo en que la conciencia, especialmente el dolor y el sufrimiento, resulta decisiva para decidir qué consideración moral deberíamos dispensar a los animales. Si estamos de acuerdo en que incluso algunos animales son capaces de sentir dolor, y en que causar sufrimiento innecesario es un error, debemos concluir que es incorrecto infligirles sufrimiento innecesario. Desarrollar más esta conciencia (decidiendo, en particular, qué ocurre si nada justifica adecuadamente el dolor que infligimos a los animales) se convierte así en una tarea moral apremiante.

«A veces se dice que los animales no hablan porque carecen de la capacidad mental para hacerlo. Esto significa: "no piensan, por eso no hablan". Pero simplemente no hablan. O mejor aún: no usan el lenguaje, exceptuando las formas más primitivas de lenguaje.»

Ludwig Wittgenstein, 1953

En el interior de la mente de los animales ¿Qué sabemos de lo que pasa en la cabeza de los animales? ¿Sienten los animales, piensan, creen? ¿Son capaces de razonar? La verdad es que sabemos muy poco sobre la conciencia de los animales. Nuestra falta de conocimiento en este ámbito representa una versión ampliada del problema de saber algo de las otras mentes *humanas* (véase la página 48). Parece que no podemos tener la seguridad de que los demás experimenten las cosas igual que nosotros, o ni siquiera de que experimenten, de modo que no es sorprendente que no estemos en mejores condiciones (posiblemente sean bastante peores) con respecto a los demás animales.

Tanto en el caso de la mente humana como en el de la del animal, todo lo que podemos hacer es usar un argumento por analogía con nuestro propio caso. Los mamíferos parecen reaccionar al dolor de un modo parecido a los humanos, pues rechazan lo que les produce dolor, prorrumpen en gritos y gemidos, etc. También en términos fisiológicos existe una uniformidad básica entre los sistemas nerviosos de los mamíferos; asimismo se han encontrado profundas similitudes en la composición genética y el origen evolutivo. Dadas todas estas similitudes, es plausible suponer que también deberían existir semejanzas en el nivel de la experiencia subjetiva. Y cuanto más próximas sean las similitudes psicológicas, y de otros aspectos relevantes, más segura será la inferencia de similitudes en la experiencia subjetiva.

El perro de Crísipo

En el mundo antiguo, no había consenso sobre si los animales pensaban y razonaban. La discusión más interesante giraba en torno al perro de Crisipo, un filósofo estoico del siglo III a. C. Se trata de la historia de un perro de caza que persigue a su presa y llega a un cruce de tres caminos; tras perder el rastro en los dos primeros caminos, el perro toma el tercero sin mayor investigación, asumiendo supuestamente el silogismo «A B o C: ni A ni B; entonces C». Este caso de lógica canina no convenció a todos los filósofos posteriores, y muchos de ellos consideraron la racionalidad como la facultad que establecía una diferencia esencial entre hombres y animales. Descartes en particular, tenía mala opinión sobre el intelecto animal, y consideraba a los animales poco más que autómatas desprovistos de cualquier forma de inteligencia La idea de que la capacidad de sufrir es la clave para admitir a los animales en la comunidad moral —el criterio que con mayor frecuencia se invoca en las discusiones recientes sobre ética animal— fue planteada por el filósofo utilitarista Jeremy Bentham: «La cuestión no es "¿razonan?" ni "¿piensan?", sino "¿sufren?"».

La experimentación con animales: ¿es correcta y sirve para algo?

La moralidad con respecto al uso de animales en la investigación médica y en las pruebas de fármacos puede considerarse de dos modos. Uno de

ellos consiste en preguntar si es correcto para nosotros usar a otros animales como meros medios para nuestros fines; si es ético causar sufrimiento a los animales (asumiendo que sufran) y quebrantar sus derechos (asumiendo que puedan tenerlos) para mejorar nuestra salud, probar fármacos, etcétera. Éste es un aspecto de la pregunta general sobre la posición moral que debemos adoptar con respecto a los animales (véase la página 108).

La otra consideración es más práctica. Probar la toxicidad de un producto con un ratón sólo tiene valor (asumiendo que sea ético hacerlo) si el ratón y el hombre son suficientemente similares en aspectos fisiológicos relevantes como para que las conclusiones puedan aplicarse a los hombres, El problema es que esta segunda consideración pragmática alienta el uso de mamíferos como los monos y otros simios, porque desde el punto de vista psicológico son más parecidos a los humanos; pero precisamente el uso de estos animales es el que nos resulta más censurable desde el punto de vista ético.

En este sentido nuestras inferencias sobre nuestros parientes cercanos, los simios y los monos, parecen apoyarse en un terreno relativamente firme; pero el terreno es menos firme con mamíferos más alejados como las ratas y los topos. La analogía es débil pero plausible en el caso de algunos vertebrados (pájaros, reptiles, anfibios y peces), y decididamente precaria cuando se trata de invertebrados (insectos, babosas, medusas). Lo cual no significa que estos animales no sientan, o no experimenten dolor, pero sí que resulta dudoso explicar lo que sienten a partir de una analogía con nuestra conciencia. La dificultad consiste en saber cómo podría apoyarse la reivindicación de que sí sienten.

La idea en síntesis: ¿crueldad con los animales?

26 ¿Tienen derechos los animales?

En todo el mundo, durante cada año de la primera mitad de la década de 2000:

se usaron aproximadamente 50 millones de animales en investigaciones científicas y pruebas farmacológicas;

se produjeron unos 250 millones de toneladas de carne;

se capturaron unos 200 millones de toneladas de pescado y otras especies acuáticas en mares y ríos.

Las cifras son aproximadas (especialmente las relativas a la investigación, la mayor parte de las cuales no se registran), pero es evidente que cada año se usa una inmensa cantidad de animales para satisfacer intereses humanos. Muchos de nosotros (cada vez más) diríamos que son «explotados» o «sacrificados» en vez de «usados». Porque muchos consideramos el uso de los animales (como alimento o como instrumento de investigación) como algo moralmente indefendible y como una violación de los derechos fundamentales de los animales.

El fundamento de los derechos de los animales ¿Cómo podemos fundamentar los derechos de los animales? Un argumento común, de carácter esencialmente utilitarista, es el siguiente:

1. los animales sufren;
2. el mundo es mejor sin sufrimiento innecesario; de modo que
3. no debe infligirse sufrimiento innecesario a los animales.

La primera premisa ha sido la más debatida recientemente (véase la página 104). Parece muy improbable suponer que animales como los simios y los monos, que tanto se parecen a nosotros en aspectos relevantes, no tengan la capacidad de sentir algo muy similar a nuestro dolor. Sin embargo, parece improbable que animales como las esponjas marinas o las medusas, que tienen sistemas nerviosos muy simples, sientan algo remotamente parecido al dolor humano. La dificultad radica entonces en dónde establecer la frontera (algo que ocurre a menudo cuando se trata de establecer las fronteras [véase la página 89]) y resulta difícil eludir un considerable tufillo de arbitrariedad. Podemos acordar un ponderado «*algunos* animales sufren», pero un inquietante signo de interrogación pende sobre cuáles son realmente.

Cronología

La segunda premisa podría parecer irrecusable (salvo a los raros masoquistas), pero una vez más entraña un peligro y es que puede llegar a resultar vacía. Es posible cuestionar la pretensión de esta premisa estableciendo una distinción entre el dolor y el sufrimiento. Se supone que el ultimo es una emoción compleja, que implica tanto el recuerdo del dolor pasado como la anticipación del dolor por venir, mientras que el dolor en sí mismo no es más que una fugaz sensación del presente; lo que cuenta cuando se trata de hacer valoraciones morales es el sufrimiento, pero los animales, o algunos animales, sólo son capaces de sentir dolor. Pero incluso si aceptamos esta distinción, parece poco razonable pretender que el dolor no es algo malo, por más que el sufrimiento sea peor.

Mucho más problemático resulta el « innecesario» de la segunda premisa. Porque no sirve para persuadir a un oponente que defienda que un poco de dolor animal es un precio aceptable por los beneficios humanos, ya que permite la mejora de la salud, la seguridad de los fármacos, etcétera. Desde el punto de vista utilitarista, el argumento aparentemente apela a una especie de cálculo del dolor, sopesando el sufrimiento animal y el beneficio humano; pero el cálculo necesario —dificilísimo incluso si sólo se tratara de medir el sufrimiento humano — parece completamente irrealizable cuando el sufrimiento animal se incorpora a la ecuación.

«**Llegará un día en que el resto de animales de la creación adquieran los derechos que la tiranía nunca debería haberles negado.**»

Jeremy Bentham, 1788

Este ataque a las premisas afecta inevitablemente a la conclusión. Por desgracia, hay que admitir que, a lo sumo, el argumento se traduce en la reivindicación de no hacer daño a algunos, tal vez muy pocos, animales, a menos que hacerlo brinde algún beneficio, tal vez mínimo, a los humanos. Desde este punto de vista « los derechos de los animales» se reducen al derecho de un pequeño número de animales a los que no debe infligirse sufrimiento, a menos que hacerlo nos proporcione algún beneficio, por pequeño que sea, a los humanos.

Especismo

La mayoría de los individuos no someterían a un igual a condiciones de hacinamiento y suciedad para terminar comiéndoselo; ni probaría

productos químicos de propiedades desconocidas con niños; ni modificaría genéticamente a humanos para estudiar su biología. ¿Existen razones para tratar de este modo a los animales? Debe existir alguna justificación moral atendible (según los defensores de los derechos de los animales) para rechazar conceder la misma consideración a los intereses de los animales que a los de los hombres. De otro modo estaríamos ante un caso de prejuicio o de intolerancia, de discriminación por cuestiones de especie, es decir, de «especismo»: una falta de respeto fundamental de la dignidad y las necesidades de los animales que no son humanos, tan ilegítima como la discriminación por cuestiones de género o de raza. ¿Es una falta evidente favorecer a nuestra propia especie? Los leones, por ejemplo, por lo general son más considerados con otros leones que con las gacelas; ¿por qué no deberían los humanos comportarse con una parcialidad similar? Se han sugerido muchas razones:

- los humanos tienen un nivel de inteligencia más elevado que los animales, al menos potencialmente;
- la depredación es natural (los animales comen otros animales);
- los animales son especialmente criados para ser comidos/usados en experimentos (si no fuera por eso no existirían);
- necesitamos comer carne, aunque millones de personas aparentemente sanas no lo necesitan;
- los animales no tienen alma (¿estamos seguros de que los humanos la tienen?),

Es difícil rebatir estas justificaciones y, en general, resulta complejo establecer criterios que abarquen a todos los humanos y excluyan a todos los animales. Por ejemplo: si decidimos que lo decisivo es la superioridad intelectual, ¿utilizaríamos este criterio para justificar el uso de niños o de personas mentalmente discapacitadas, con un nivel de inteligencia menor que un chimpancé, en un experimento científico? O si decidimos que lo decisivo es la «naturalidad», advertiremos de inmediato que hay muchas cosas que los animales, incluidos los humanos, hacen con «naturalidad» que no nos gustaría en absoluto promover: algunos leones macho, llevados por su naturaleza, matan a las crías de sus rivales, pero semejante comportamiento estaría muy mal visto entre los humanos.

Las tres erres

El intenso debate sobre el bienestar y los derechos de los animales se centra en dos preguntas: ¿deberían usarse animales en experimentos y (si lo fueran) cómo deberían ser tratados en práctica? Como resultado de estas preguntas, se aceptan actualmente tres principios generales, « las tres erres» , como guía para la técnica experimental humana:

- **reemplazar** a los animales siempre que sea posible por otras alternativas;
- **reducir** el número de animales que se usan en experimentos hasta un nivel la coherente con la producción adecuada de datos estadísticos,
- **refinar** las técnicas experimentales para reducir o erradicar el sufrimiento de los animales.

¿Están bien los derechos? No es una conclusión que pueda complacer a ningún defensor serio de los derechos de los animales. Pero ha habido justificaciones más sólidas y sofisticadas que la versión resumida antes, y todas ellas pretenden ofrecer una versión menos débil del tipo de derechos de los que deberían disfrutar los animales. Peter Singer ha sido el defensor de la concepción utilitarista sobre el asunto, y otra línea muy influyente ha sido la deontológica, defendida por el norteamericano Torn Reagan. Según Regan, los animales son « sujetos de una vida» (al menos los animales con un cierto grado de complejidad); este hecho les confiere algunos derechos básicos, que son los que se violan cuando un animal es tratado como un simple pedazo de carne o como un representante de los hombres en las pruebas con fármacos. En este sentido, los animales no deberían ser sometidos a análisis de costes-beneficios que pueden ser muy perjudiciales para una concepción utilitarista.

Existen considerables dificultades para defender una concepción en que los derechos de los animales sean equiparables a los derechos humanos, e incluso algunos autores se preguntan si es apropiado y útil emplear la noción de derechos para el caso. Se suele suponer que los derechos imponen deberes u obligaciones a sus titulares; hablar de derechos implica algún tipo de reciprocidad (algo que precisamente nunca existirá entre los animales y los hombres). Lo que está en juego es un problema real —el trato adecuado y humano que debemos a los animales— que se oscurece al plantearlo provocadoramente en el lenguaje de los derechos.

La idea en síntesis: ¿se equivocan los humanos?

27 Las formas del razonamiento

Los argumentos son los ladrillos con que se construyen las teorías filosóficas; la lógica es la argamasa que mantiene unidos los ladrillos. Las buenas ideas son más bien poca cosa a menos que se apoyen en buenos argumentos: deben estar justificadas racionalmente, y ello no se consigue sin unos fundamentos lógicos firmes y rigurosos. Los razonamientos expuestos con claridad son susceptibles de evaluación y crítica, y este proceso continuo de reacción, revisión y rechazo es el que permite el progreso filosófico.

Un razonamiento es un movimiento racional autorizado que va desde unos hechos aceptados (*las premisas*) hasta un punto que debe ser probado o demostrado (*la conclusión*). Las premisas son las proposiciones básicas que deben aceptarse, al menos provisionalmente, para que el razonamiento pueda desarrollarse. Las premisas mismas pueden establecerse de muchos modos, a partir de la lógica, o sobre la base de la evidencia, es decir, empíricamente, o también pueden ser conclusiones de razonamientos anteriores; pero en cualquier caso deben sostenerse independientemente de la conclusión para evitar la circularidad. El movimiento desde las premisas hasta la conclusión es una inferencia, la firmeza de la cual determina la resistencia del razonamiento. Distinguir las inferencias correctas de las incorrectas es la tarea central de la lógica.

El papel de la lógica La lógica es la ciencia de analizar los razonamientos y de establecer principios o fundamentos sobre los que puedan sostenerse inferencias sólidas. Su preocupación no son los contenidos particulares de los razonamientos sino su estructura general y su forma. Así, « todos los pájaros tienen plumas; el petirrojo es un pájaro; de modo que el petirrojo tiene plumas» , los lógicos abstraen la forma así: « Todos los P son Q; *a* es P; luego *a* es Q» , fórmula en la que los términos particulares se sustituyen por símbolos, y la firmeza de la inferencia puede determinarse con independencia del asunto en cuestión. El estudio de la lógica se ceñía antiguamente a inferencias simples de este tipo (*silogismos*), pero desde principios del siglo XX se ha transformado en una herramienta analítica inmensamente sofisticada y sutil.

Cronología



La lógica aristotélica y matemática

Hasta finales del siglo XIX la ciencia de la lógica discurrió con pocos

cambios, muy próxima a la manera en que Aristóteles la había planteado unos 2 000 años antes. Se consideraba que el modelo del razonamiento adecuado era el Gottlob silogismo, una inferencia resultante de tres proposiciones (dos premisas y una conclusión), la más célebre de las cuales es: «todos los hombres son mortales, los griegos son hombres; luego los griegos son mortales». Los silogismos fueron clasificados exhaustivamente de acuerdo con la «forma» y la «figura» a partir de las cuales era posible distinguir los válidos de los no válidos. Las limitaciones de la lógica tradicional fueron expuestas de un modo contundente en la obra de un matemático alemán, Frege, que introdujo nociones como los cuantificadores y las variables, a las que se debe la ampliación del alcance y del poder de la lógica moderna matemática (así llamada porque, a diferencia de la lógica tradicional, permite representar un razonamiento matemático).

La deducción El ejemplo que se ha ofrecido antes («todos los pájaros tienen plumas…») es un razonamiento *deductivo*. En este caso la conclusión se sigue de (está implicada en) las premisas, y el razonamiento se considera «válido». Si las premisas de un razonamiento válido son verdaderas, la conclusión es con toda seguridad verdadera, y el argumento se considera «fundado». La conclusión de un razonamiento deductivo está implícita en las premisas; dicho de otro modo, la conclusión no «va más allá» de sus premisas, ni dice nada que no estuviera ya implicado en ellas. Otro modo de expresar esto, que revela el carácter lógico que subyace al razonamiento, es señalar que no es posible aceptar las premisas y rechazar la conclusión sin incurrir en contradicción.

1670 Fe y razón **1739** Ciencia y pseudociencia **1905** El rey de Francia es calvo

¿Paradoja o falacia?

«Al prisionero lo colgarán el próximo sábado a lo sumo, y no sabrá por anticipado qué día se ejecutará la sentencia.» Parece espantoso, pero el astuto prisionero se reconforta pensando: «No pueden colgarme el sábado, porque en ese caso el viernes lo sabría por anticipado. De modo que el último día posible es el viernes. Pero tampoco puede serlo, porque lo sabría el jueves…». Y así fue retrocediendo día a día hasta el presente, y le alivió descubrir que la ejecución no tendría lugar jamás. Así que el prisionero se llevó un pequeño sobresalto cuando fue colgado

finalmente el siguiente martes.

¿Paradoja o falacia? Tal vez ambas cosas. El relato, conocido como la paradoja de la predicción, es paradójico porque una línea de razonamiento aparentemente impecable desemboca en una conclusión manifiestamente falsa, tal como acaba descubriendo el compungido prisionero. Las paradojas incorporan argumentos de este tipo que conducen a conclusiones aparentemente contradictorias o inaceptables. A veces resulta imposible sacar una conclusión, y ello obliga a examinar de nuevo las creencias o supuestos en los que se apoya el razonamiento; o puede ocurrir que se haya colado alguna falacia, un error de razonamiento, en el mismo argumento. En cualquier caso, las paradojas merecen atención filosófica, porque señalan invariablemente confusiones o incoherencias en nuestros conceptos y razonamientos.

Algunas de las paradojas más célebres (muchas de las cuales se analizarán en las próximas páginas) se resisten tenazmente a ser resueltas, y siguen dejando perplejos a los filósofos.

La inducción El otro procedimiento principal para pasar de las premisas a la conclusión es la inducción. En un razonamiento característicamente *inductivo*, una ley o principio general se infiere de observaciones particulares sobre cómo son las cosas en el mundo. Por ejemplo, a partir de un número de casos en los que se observa que los mamíferos paren crías vivas, se infiere, inductivamente, que todos los mamíferos lo hacen. Un razonamiento semejante no puede ser nunca válido (en el sentido en el que un razonamiento deductivo puede ser válido) en la medida en que su conclusión no se sigue *necesariamente* de sus premisas; es decir, es posible que las premisas sean verdaderas pero la conclusión falsa (como ocurre en el ejemplo ofrecido, donde la conclusión es falsa dada la existencia de mamíferos como el ornitorrinco, que pone huevos). Esto ocurre porque el razonamiento inductivo siempre va más allá de las premisas, que nunca *implican* una conclusión dada, sino tan sólo la apoyan o la hacen posible hasta cierto punto. De modo, pues, que los razonamientos inductivos son generalizaciones o extrapolaciones de varios tipos: de lo particular a lo general; de lo observado a lo no observado; de acontecimientos o estados de hecho pasados, o presentes, a futuros.

El razonamiento inductivo es omnipresente e indispensable. Resultaría imposible prescindir en nuestras vidas cotidianas de la observación de regularidades y continuidades en el pasado y en el

«**Por el contrario —prosiguió Tweedledee— si fuera así, podría ser; y si era así, debería ser: pero**

presente, pues ellas nos permiten hacer predicciones sobre cómo serán las cosas en el futuro. Asimismo, las leyes y supuestos de la ciencia a menudo consisten en casos paradigmáticos de inducción (véase la página 136). Pero ¿es justificable extraer tales inferencias? El filósofo escocés David Hume pensaba que no existe una base *racional* para nuestra confianza en las inducciones. Sostenía que los razonamientos inductivos presuponen una creencia en la «regularidad de la naturaleza», y a partir de ésta asume que el futuro será similar al pasado cuando se den condiciones similares relevantes. Pero ¿qué razones podrían existir para asumir semejante cosa, aparte de razones inductivas? Si la supuesta regularidad de la naturaleza sólo puede justificarse inductivamente, no es posible —sin incurrir en un argumento circular— usar el procedimiento en defensa de la inducción. De un modo parecido, hay quienes han intentado justificar la inducción sobre la base de los sucesos pasados: efectivamente *funciona*. Pero la suposición de que seguirá funcionando en el futuro sólo puede inferirse *inductivamente* a partir de los sucesos pasados, de modo que el argumento no consigue salir del atolladero. Desde el punto de vista de Hume no podemos hacer otra cosa que razonar inductivamente (y ni siquiera sugiere que no debiéramos razonar de este modo), pero insiste en que procedemos así por costumbre y hábito aunque no esté racionalmente justificado. El llamado «problema de la inducción» que Hume dejó en herencia sigue siendo un asunto de debate activo en la actualidad, especialmente en la medida en que afecta a los fundamentos de la ciencia.

como no lo es, no. Pura lógica.»

Lewis Carroll, 1871

La idea en síntesis: ¿existen razonamientos infalibles?

28 La paradoja del barbero

En un pueblo vive un barbero que afeita única y exclusivamente a los que no se afeitan ellos mismos. ¿Quién afeita al barbero, pues? Si se afeita él mismo, no se afeita; si no se afeita él mismo, se afeita.

> **«La filosofía consiste en empezar con algo tan simple que parezca irrelevante, y en terminar con algo tan paradójico que nadie pueda creerlo.»**
>
> **Bertrand Russell, 1918**

A primera vista, el rompecabezas que encierra la paradoja del barbero no parece tan complicado de entender. Se trata de una situación que en principio parece plausible, pero que desemboca de forma inevitable en una contradicción. La descripción en apariencia inocente del trabajo (un hombre «que afeita única y exclusivamente a los que no se afeitan ellos mismos») es, de hecho, lógicamente imposible, pues el barbero no puede pertenecer ni al grupo de los que se afeitan ellos mismos, ni al de los que no lo hacen, sin contradecir la descripción de sí mismo. Un hombre que se adecuara a la descripción del barbero no puede existir (lógicamente). De modo que no existe tal barbero: problema resuelto.

El significado de la paradoja del barbero no se cifra en su contenido sino en su forma. Desde el punto de vista estructural, esta paradoja es similar a otro problema más importante conocido como la paradoja de Russell, que trata, no ya del afeitado de la gente de un pueblo sino de conjuntos matemáticos y de sus contenidos. Tal paradoja es bastante más difícil de resolver: no es exagerado decir que, un siglo atrás, fue la principal responsable de la crisis de los fundamentos básicos de las matemáticas.

Cronología

c. 350 a. C.	*c.* 300 a. C.
Las formas del razonamiento	La paradoja sorites

«Esta proposición es falsa»

El problema de la autorreferencia que encierra la paradoja del barbero y la paradoja de Russell es común a un buen número de rompecabezas filosóficos muy célebres. Tal vez, el más conocido sea la llamada «paradoja del mentiroso», los orígenes de la cual, supuestamente, se remontan al siglo VII a. C., cuando parece que un griego llamado Epiménides —de Creta— dijo que «todos los ciudadanos de Creta eran unos mentirosos». La versión más simple es la proposición «Esta proposición es falsa», que es falsa si es verdadera y verdadera si es

falsa. La paradoja puede captarse en un par de proposiciones: en un lado de un pedazo de papel leemos «La proposición que se encuentra en el otro lado del papel es falsa»; y en el otro lado del papel «La proposición que se encuentra en el otro lado del papel es verdadera». En este planteamiento cada una de las proposiciones es en sí misma inobjetable, de modo que resulta difícil desechar la paradoja como si fuera un simple absurdo, como pretenden algunos.

Otra variante interesante es la paradoja de Grelling. En este caso, están implicadas las nociones de «palabras autológicas» (palabras que ilustran lo que describen), por ejemplo la palabra «esdrújula», que es ella misma esdrújula; y de «palabras heterológicas» (palabras que no ilustran lo que describen), por ejemplo «largo», que es una palabra corta. Toda palabra debe ser de un tipo u otro, de modo que consideremos lo siguiente: ¿la palabra «heterológico» es heterológica? Si lo es, no lo es; si no lo es, lo es. Todo parece indicar que no hay modo de salir de la barbería.

Russell y la teoría de conjuntos La idea de los conjuntos es fundamental para las matemáticas, pues constituyen los objetos más puros que pueden examinarse. El método matemático implica definir grupos (conjuntos) de elementos que satisfacen determinado criterio, como el conjunto de todos los números mayores que 1 o el de los números primos; se realizan entonces operaciones de tal modo que luego puedan deducirse propiedades sobre los elementos incluidos en el conjunto o los conjuntos en cuestión. Desde una perspectiva filosófica, los conjuntos tienen un interés particular, pues la idea de que todo lo relativo a las matemáticas (los números, las relaciones, las funciones) podría formularse exhaustivamente en el interior de una teoría de conjuntos dio alas a la ambición de usar los conjuntos para fundamentar las matemáticas sobre bases puramente lógicas.

1901 La paradoja del barbero

1905 El rey de Francia es calvo

Claridad de pensamiento

Los razonamientos filosóficos a menudo son complejos y deben expresarse con gran precisión. Pero a veces los filósofos se dejan llevar un poco por la solemnidad de su propio intelecto, y entonces seguir sus argumentos resulta tan complicado como intentar digerir una piedra. Si

entender el fuera de juego en fútbol te pareció complicado, intenta seguir el razonamiento de Bertrand Russell al definir «el número de una clase», a ver qué te parece:

«Este método consiste en definir el número de una clase como la clase de todas las clases similar a la clase dada. Un miembro de esta clase de clases (considerado como predicado) es una propiedad común de todas las clases similares y de ninguna otra; además, cada clase del conjunto de clases similares tiene con el conjunto una relación que no tiene con nada más, y que cada clase tiene con su propio conjunto. Así, se satisfacen completamente las condiciones de esta clase de clases, y su mérito consiste en estar determinada cuando se da una clase, y en ser distinta para dos clases que no son similares. He aquí, pues, una definición irreprochable del número de una clase en términos puramente lógicos».

A principios del siglo XX, el matemático alemán Gottlob Frege intentaba definir la totalidad de la aritmética en términos lógicos por medio de la teoría de conjuntos. En la época se asumía que no había restricciones en las condiciones que podían usarse para definir conjuntos. El problema, que reconoció el filósofo Bertrand Russell en 1901, se centraba en la cuestión de los conjuntos que están comprendidos en su propio conjunto. Algunos conjuntos se tienen a sí mismos como elemento: por ejemplo, el conjunto de objetos matemáticos es él mismo un objeto matemático. Otros no: el conjunto de números primos no es él mismo un número primo. Consideremos ahora el conjunto de todos los conjuntos que no están comprendidos en su propio conjunto. ¿Forma parte este conjunto de su propio conjunto? Si forma parte, no forma parte: y si no forma parte, forma parte. Es decir, ser un elemento de este conjunto depende de no ser un elemento del conjunto.

«**Existen pocas cosas más indeseables para un científico que ver hundirse los fundamentos en el momento en que uno daba por acabado el trabajo. Ésta es la situación en la que me puso una carta del señor Bertrand Russell cuando estaba a punto de dar por terminado mi trabajo.**»

Gottlob Frege, 1903

Una perfecta contradicción, y de ahí la paradoja (similar a la del barbero). Sin embargo, a diferencia de lo que ocurre en el caso del barbero, no es posible deshacerse del temible conjunto (por lo menos no es posible hacerlo sin socavar

la teoría de conjuntos tal como se entendía entonces).

La existencia de contradicciones en el seno de la teoría de conjuntos, que mostró la paradoja de Russell, puso en evidencia que la definición y los procedimientos matemáticos de los conjuntos eran fundamentalmente erróneos. Puesto que cualquier proposición puede probarse, lógicamente, a partir de la contradicción, ello implicaba, lamentablemente, que toda y cada una de las pruebas, aunque no necesariamente inválidas, podían no ser probadas como válidas. De modo que las matemáticas debían reconstruirse desde los fundamentos. La clave de la solución se encontraba en la incorporación de restricciones adecuadas en los principios que rigen la inscripción de un elemento en un conjunto. Russell no sólo expuso el problema sino que fue uno de los primeros que buscó una solución y, aunque su tentativa fue sólo parcialmente eficaz, contribuyó a que sus sucesores tomaran el camino adecuado.

La idea en síntesis: si es, no es, si no es, es

29 La falacia del jugador

«Con la cabeza a mil por hora, Monte y Cario escrutaban al crupier que empezaba a recoger las fichas perdidas. No habían apostado en los últimos juegos, pues preferían oler antes el juego de la mesa. Pero la impaciencia crecía a medida que los números rojos salían una y otra vez, ya iba a ser la quinta; no aguantaban más.
—Si no te metes seguro que no ganas —pensaron los dos, no eran muy originales…

» Cuando se disponían a jugar sus fichas, Monte pensó:

» —¡Cinco veces seguidas el rojo! Seis es imposible. ¿Cuántas probabilidades hay contra el rojo? Por la ley de la probabilidad tiene que salir el negro.

» En el mismo momento, Cario pensaba:

» —¡Caramba, los rojos están que arden! Esta vez no se me escapan. Tiene que salir el rojo.

» —*Ríen ne va plus*… No hay más apuestas —anunció el crupier.»

¿Quién es más probable que gane, Monte o Cario? La respuesta tal vez sea que ganará Cario, y todavía más probable, ninguno. Lo cierto es que ambos (quizás igual que millones de personas reales a lo largo de la historia, pues se han encontrado dados del año 2750 a. C. aproximadamente) incurren en la llamada falacia del jugador (o de Monte Cario).

Cronología

c. 2750 a. C.
La falacia del jugador

«Tiene que salir el negro» Monte tiene razón en que cinco veces seguidas el rojo es inusual: la probabilidad (en una mesa donde el juego sea limpio, y al margen de las apuestas del cero o del doble cero; véase el cuadro) es de 1 contra 32; y las probabilidades de seis veces seguidas son incluso más bajas: de 1 contra 64. Pero estas probabilidades sólo se aplican al principio de la secuencia, antes de que la ruleta haya empezado a girar. El problema para Monte es que la

El error de razonamiento que se observa en la falacia del jugador lo ilustra maravillosamente la historia del hombre al que pillaron en un avión con una bomba como equipaje de mano.

« Las probabilidades de que haya una bomba en un avión son poquísimas —explicó a la policía—. ¡Así que imaginen lo

posibilidad relativamente rara (cinco veces seguidas el rojo) *ya se ha producido* sin tener nada que ver en absoluto con que el siguiente número sea el rojo; la probabilidad de que eso ocurra es, como siempre, de 1 contra 2 o de 50/50. Las ruletas —como las monedas, los dados, o las bolas de la lotería— no tienen memoria, de modo que no pueden tener en cuenta lo que ha ocurrido en el pasado con vistas a compensar, o incluso a propiciar, ciertas cosas: la improbabilidad de cualquier hecho pasado o de cualquier secuencia de hechos (siempre que sean azarosos e independientes) no tiene nada que ver con la probabilidad de un hecho futuro. Suponer que es de otro modo conduce a la falacia del jugador.

improbable que es que hubiera dos!»

La casa siempre gana

Los juegos de casino siempre tienen algún tipo de « margen de la casa» , lo cual significa que las probabilidades están levemente inclinadas a favor de la banca. Por ejemplo, en la ruleta, existe una casilla (en Europa el cero) o dos (en Estados Unidos, el cero y el doble cero) que no son ni rojas ni negras, de modo que las posibilidades de ganar al rojo o al negro son un poco inferiores a la mitad. Asimismo, en el 21 de naipes (*vingt-et-un*), hay que *batir* a la banca: si la banca obtiene un *Black Jack* (21 tantos con dos cartas) bate al jugador. Y aunque un jugador puede ganar a la casa, casi siempre, a lo largo de todos los tiempos, la casa suele ganar más juegos.

1950
El dilema del prisionero

1963
La teoría tripartita del conocimiento

«**Me siento como un fugitivo de la ley estadística.**»

Bill Mauldin, 1945

«¡El rojo está que arde!» ¿Acaso Cario procede mejor? Probablemente no. Como Monte, intenta predecir un resultado futuro a partir de hechos que en apariencia no tienen nada que ver. Y si los hechos a los que asiste son azarosos, también él está incurriendo en la falacia del jugador. Esta falacia sólo está asociada a resultados que son genuinamente independientes. Si un caballo, por ejemplo, gana cuatro carreras seguidas, podríamos considerar que existen suficientes evidencias de que puede ganar una quinta. Pero si una ficha gana 20

veces seguidas, cabe suponer razonablemente que la ficha está trucada en vez de pensar que semejante cosa improbable ocurra por puro azar.

Si juegas a la lotería, espera sentado…

¿Cuántas probabilidades tienes de que salgan tus seis números dos veces seguidas en la lotería nacional inglesa?

Aproximadamente 1 contra 200 000 000 000 000 (200 millones de millones).

No muchas, de modo que tendrías que ser un auténtico idiota para escoger el mismo número que la semana pasada… Puede que sí, pero no más idiota que si escoges seis números distintos. Se trata de otro caso de la falacia del jugador: una vez *ya ha salido* una serie de números el hecho de que vuelva a salir no constituye una selección mejor o peor (14 millones contra 1 no es mucho más tentador).

De modo que quienes se preguntan si la mejor estrategia es jugar siempre con la misma serie de números o cambiar cada semana deberían saber que no existe demasiada diferencia (es mucho mejor cavar en el jardín a ver si hay algún tesoro enterrado).

Las leyes de la probabilidad

Las leyes de la probabilidad se invocan a menudo para refrendar el pensamiento falaz de los jugadores.

Lo que pretenden, en resumen, es que es más probable que ocurra algo en el futuro si ha ocurrido con menos frecuencia de la esperada en el pasado

(o, a la inversa, que es menos probable que ocurra en el futuro porque ha ocurrido con mucha frecuencia en el pasado).

A partir de ahí se supone que las cosas quedarán «fuera de toda posibilidad a la larga».

El atractivo de esta ley falaz se debe, en parte, a su similitud con una auténtica ley estadística: la ley de las cifras elevadas. De acuerdo con esta ley, si lanzamos una ficha que no esté trucada unas cuantas veces, por ejemplo 10, la posibilidad de que gane disminuye considerablemente con respecto de la media (estadística), que es de 5, pero si la lanzamos un

número elevado de veces, por ejemplo 1 000, la posibilidad de que gane se acerca más al promedio (500). Y cuanto mayor sea el número de tiradas, más se acercará al promedio. De modo que, en una serie de hechos azarosos de similar probabilidad, es cierto que si la serie se extiende lo suficiente hay cosas que podrían incluso quedar descartadas.

Sin embargo, esta ley estadística no tiene nada que ver con la posibilidad de que ocurra algún hecho particular: el hecho que acontece ahora no tiene memoria de ninguna desviación de la media, y no puede alterar su curso para corregir un desequilibrio anterior.

No hay consuelo para el jugador.

Del mismo modo, una secuencia de cuatro rojos seguidos *podría* indicar que la ruleta está trucada o amañada. Sin embargo, aunque no sea imposible, cuatro rojos seguidos es muy inusual y del todo insuficiente por sí mismo para garantizar cualquier conclusión. En ausencia de cualquier otra evidencia, Cario es tan iluso como Monte.

La idea en síntesis: contra la probabilidad

30 La paradoja sorites

Supon (si en tu caso no es cierto) que tienes la cabeza completamente llena de pelos. Eso quiere decir que es probable que tengas unos 100 000 pelos. Arráncate uno. ¿Acaso te has convertido en un calvo? Claro que no. Un solo pelo no cambia nada. Con 99 999 pelos sigues teniendo una estupenda melena.

Todos nosotros coincidiríamos, excepto si somos calvos, en que perder un pelo no nos hace ser calvos. Pero si nos arrancamos otro, y otro, y otro… Al final, si insistimos lo suficiente, no nos quedará ni uno y seremos perfectamente calvos. De modo que aparentemente pasamos de un estado de indudable no calvicie a un estado de indudable calvicie merced a una serie de pasos que individualmente nunca tendrían ningún efecto. Pero ¿cuándo se produce el cambio de un estado a otro?

Ésta es una versión de un problema célebre, atribuido habitualmente al lógico griego de la Antigüedad Eubúlides de Mileto, y que se conoce como la paradoja sorites. « Sorites» viene de la palabra griega « soros» que significa « montón» , pues la formulación original de este rompecabezas tenía por objeto un montoncito de arena. Planteado como una suma (de granos de arena) en vez de como una sustracción (de pelos), el razonamiento se presenta del siguiente modo:

1 grano de arena no hace un montón.

Si 1 grano de arena no hace un montón, 2 tampoco.

Si 2 granos no hacen un montón, 3 tampoco.

(*y así hasta*…)

Si 99 999 granos no hacen un montón, 100 000 tampoco.

Así pues, 100 000 granos de arena no hacen un montón.

Cronología

c. 350 a. C.	*c.* 300 a. C.
Formas de razonamiento	La paradoja sorites

Pero nadie estaría dispuesto a aceptar esta conclusión. ¿Qué puede haber fallado?

Problemas de vaguedad Frente a una conclusión inadmisible como ésta, es necesario hacer marcha atrás en el razonamiento que nos ha llevado a ella. Debe haber algo erróneo en las premisas en las que se basa el razonamiento o algún error en el desarrollo. De hecho, a pesar de ser un rompecabezas tan antiguo,

todavía no se ha llegado a un consenso claro sobre cómo hacer frente a esta paradoja, y existen diversas propuestas.

Una salida consiste en insistir, como se ha hecho, en que existe un punto en el que añadir un grano produce una diferencia; en que hay un número preciso de pelos que marca el límite entre un calvo y uno que no lo es. Si existe tal frontera, es evidente que no sabemos dónde se encuentra, y cualquier línea divisoria que se proponga parece inevitablemente arbitraria: ¿1 001 granos, por ejemplo, harían un montón, pero 999 no? Esto es toda una ofensa al sentido común y a nuestras intuiciones habituales.

Parece mejor solución analizar con atención un supuesto importante que subyace al argumento: la idea de que el proceso de construcción mediante el que alguien que no es calvo se queda calvo puede analizarse completamente, y reducirse a una serie de distintas sumas de granos. Es evidente que hay varios pasos distintos, pero no lo es menos que ninguno de ellos es plenamente constitutivo del proceso general de hacer un montón.

Lógica terminal

Los fumadores, dada su tendencia a esconder la cabeza bajo el ala, suelen ser susceptibles al tipo de razonamiento fallido que subyace a la paradoja sorites. El fumador razona, de un modo plausible, que «uno más no va a matarme». Una vez ha establecido esto, cae sin esfuerzo en una progresión sorítica que le permite decirse, «uno más después del siguiente no va a matarme». Y así sigue, aunque no hasta el infinito. La probable verdad de que un solo cigarrillo no va a matarte (aunque es más que probable que la suma de cigarrillos fumados sí lo consiga) representa una victoria pírrica para el fumador veterano.

1901	1905	1953
La paradoja del barbero	El rey de Francia es calvo	El escarabajo en la caja

Lógica difusa

La lógica tradicional es bivalente: ello significa que sólo caben dos valores de verdad, que cada proposición debe ser o bien falsa o bien verdadera. Pero la vaguedad inherente a muchos términos, y que atestigua la paradoja sorites, sugiere que este requisito es demasiado rígido si la lógica debe dar cuenta de toda la amplitud y de toda la

complejidad del lenguaje natural.

La lógica difusa, iniciada por el matemático Lofti Zadeh, se desarrolló para dar cuenta de las imprecisiones y de los distintos grados de verdad. La verdad se presenta, entonces, como un continuo entre lo verdadero (1) y lo falso (0). Así, por ejemplo, una proposición particular que es «parcialmente verdadera» o «más o menos verdadera» puede representarse como verdad en un grado 0,8 y como falsedad en un grado 0,2. La lógica difusa ha sido particularmente importante para la investigación en IA (inteligencia artificial), donde sistemas de control «inteligentes» deben poder responder a imprecisiones o matices del lenguaje natural.

Este análisis fallido pasa por alto que la transición entre no ser un montón y serlo es un continuo, de modo, pues, que no existe un punto preciso en el que pueda decirse que tiene lugar el cambio (sobre problemas similares relacionados con la vaguedad, véase la página 91). A su vez, esto nos muestra algo acerca de la clase general de casos a los que puede aplicarse la paradoja sorites: no sólo atañe a los montones de arena o a las cabezas peladas, sino también a la altura, al tamaño, a la riqueza, a la gordura y a muchísimas otras cosas. Todos éstos son esencialmente casos de vaguedad: no existe una línea divisoria clara que establezca el punto en el que se produce la diferencia con respecto a sus opuestos (la pequeñez, la pobreza, la delgadez, etc.).

«A medida que aumenta la complejidad, las proposiciones precisas pierden significado y las proposiciones sin significado pierden precisión.»

Lofti Zadeh, 1965

«No hay verdades absolutas; todas las verdades son verdades a medias. Lo diabólico es intentar tratar las verdades como si fueran absolutas.»

Alfred North Whitehead, 1953

Una consecuencia importante de todo ello es que siempre hay casos limítrofes en los que no está claro qué términos cabe aplicar. Por ejemplo, si bien es cierto que existen individuos que son claramente calvos, y otros que no lo son claramente, hay muchos otros en medio que, dependiendo del contexto y de las circunstancias, podrían considerarse de un modo u otro. Esta vaguedad inherente

significa que no siempre es adecuado decir que una proposición como « X es calvo» es inequívocamente verdadera o falsa; más bien existen grados de verdad. Esto crea una tensión entre estos términos vagos del lenguaje natural y la lógica clásica, que es *bivalente* (cada proposición debe ser o bien verdadera o bien falsa).

El concepto de vaguedad sugiere que debe revisarse la lógica clásica para captar completamente los matices del lenguaje natural. Por ello, se ha producido una corriente que desarrolla una lógica de casos confusos y de casos en los que se dan más de dos valores de verdad (véase el cuadro de la página anterior).

La idea en síntesis: ¿cuantos granos hacen falta para tener un montón?

31 El rey de Francia es calvo

Supongamos que digo «el rey de Francia es calvo». Pensaríais que me he vuelto loco, o tal vez sólo que estoy mal informado. Pero ¿es realmente falso lo que digo? Si lo es, ello significa que, de acuerdo con una ley lógica, lo contrario —«el rey de Francia no es calvo»— es verdad, aunque no suena mucho mejor. Quizá sea que ninguna de estas dos proposiciones es verdadera o falsa: las dos son perfectos despropósitos. No obstante, aunque puedan parecer frases un poco extrañas, no parecen desprovistas de significado.

¿A los filósofos les preocupan de verdad este tipo de problemas? Un caso triste, debes estar pensando: se inventan un picor para tener donde rascarse. Pues sí: en los últimos cien años la mayor parte de las energías de los cerebros filosóficos más dotados se han consagrado al rey de Francia, a pesar de que el país sea una república desde hace más de dos siglos. La preocupación en tomo a este atolladero y otros parecidos inspiró la teoría de las descripciones del filósofo inglés Bertrand Russell, propuesta por primera vez en un artículo de 1905 titulado «Sobre la denotación». Esta teoría, que se inscribía en una amplia corriente de investigaciones a cargo de filósofos anglosajones de principios del siglo XX, se fundaba en la creencia de que el análisis meticuloso del lenguaje y de su lógica subyacente era el camino seguro —tal vez el único— para conocer el mundo que podía describirse usando ese lenguaje.

Un asunto espinoso El principal interés de la teoría de las descripciones de Russell es una categoría de términos lingüísticos conocidos como descripciones definidas: «el primer hombre en la Luna»; «el número primo más pequeño»; «la montaña más alta del mundo»; «la actual reina de Inglaterra». En cuanto a su forma gramatical, el tipo de oración en el que se encuentran estas frases —por ejemplo, «El primer hombre en la Luna era norteamericano»— es similar a las llamadas «oraciones copulativas», como «Neil Armstrong era norteamericano». En este ejemplo, «Neil Armstrong» es un nombre propio, que es referencial en el sentido de que refiere a, o denota, un objeto específico (en este caso a un hombre concreto), al que se le atribuye una propiedad (en este caso la propiedad de ser norteamericano). A pesar de que algunas descripciones tienen apariencia superficial de nombres propios, existen muchos problemas que se deben a que se definen como si fueran expresiones que refirieran a algún objeto del mundo sin serlo. Ofrecer soluciones a este tipo de casos fue una de las principales motivaciones del artículo de Bertrand Russell en 1905.

Cronología

«Así "el padre de Carlos II fue ejecutado" se convierte en: "No siempre es falso de *x* que *x* engendrase a Carlos II y que *x* fuera ejecutado", y que "si *y* engendró a Carlos II, *y* es idéntico a *x*" siempre es verdadero de *y*.»

Bertrand Russell, 1905

Dos de los problemas centrales que planteaba Russell eran los siguientes:

1. Predicados informativos de la identidad

Es obvio, si lo piensas…

Si A y B son idénticos, cualquier propiedad de A es una propiedad de B, y A puede sustituirse por B en cualquier oración que la contenga sin alterar su verdad o falsedad. Pues bien: Jorge IV quería saber si Scott era el autor de *Waverley*. Puesto que Scott era efectivamente el autor de esta novela, podemos sustituir « Scott» por « el autor de *Waverley*» y descubrir que Jorge IV quería saber si Scott era Scott. Pero no parece que eso fuera lo que quería saber. « Scott es el autor de *Waverley*» es una oración informativa mientras que « Scott es Scott» no lo es.

La angustia existencial

Muchas descripciones definidas no denotan *nada*. Así, por ejemplo, podríamos querer decir: « El mayor número primo *no existe*» . Pero es evidentemente absurdo decir que algo no existe. Es como decir que algo que existe no existe: una pura contradicción. El análisis que hace Russell de tales proposiciones explica que estas expresiones no denotativas carecen de significado a menos que aceptemos asumir la inoportuna carga metafísica de las entidades no existentes. Naturalmente, la parte más discutible de esta —posible— carga es Dios; los defectos evidentes de uno de los argumentos más importantes de la existencia de Dios (el argumento ontológico; véase la página 164) se encuentran expuestos en el análisis de Russell.

2. Preservar las leyes de la lógica

De acuerdo con la *ley del tercero excluido* (una ley de la lógica clásica) si « A es B» es falso, « A no es B» debe ser cierto. Así pues, si la proposición « el rey de Francia es calvo» es falsa (como parece ser en caso de pronunciarse en el siglo XXI), « el rey de Francia no es calvo» debe ser verdadera. Pero resulta que también es falsa. Si una proposición y su negación son falsas, la lógica parece fatalmente menoscabada.

La solución de Russell La solución a cada uno de estos problemas, desde el punto de vista de Russell, consiste simplemente en dejar de tratar las descripciones definidas en cuestión como si refirieran expresiones ocultas. Las apariencias en estos casos son engañosas: aunque las distintas proposiciones del ejemplo que se han citado antes tienen la forma gramatical de oraciones copulativas, no tienen su forma lógica; y es la estructura lógica la que debe determinar si las proposiciones son verdaderas o falsas, y si justifican las inferencias que puedan derivarse de ellas.

«Si enumeramos las cosas que son calvas, y luego aquellas que no son calvas, no encontraremos al rey de Francia en ninguna de las dos listas. Los hegelianos, que adoran las síntesis, posiblemente concluirían que usa peluca.»

Bertrand Russell, 1905

Dejando atrás el modelo referencial de sujeto-predicado, Russell propone como alternativa que las proposiciones que contengan descripciones definidas sean tratadas como proposiciones de « cuantificación existencial» . Así, de acuerdo con su análisis, una proposición con la forma general « F es G» puede descomponerse en tres afirmaciones: « Existe un F» ; « sólo una cosa es F» ; « si algo es F, entonces es G» . Gracias a este tipo de análisis, Russell disipa rápidamente los diversos misterios en torno a las cabezas coronadas de Europa:

1. « Scott es el autor de *Waverley*» se analiza como: « Existe un ente, y sólo un ente, que es el autor de *Waverley*, y ese ente es Scott» . Es evidente que una cosa es que Jorge IV se pregunte si esto es verdad; y otra muy distinta es que se pregunte por el principio de identidad que implicaba la oración en el modelo referencial.

2. « El actual rey de Francia es calvo» se convierte en el análisis de Russell en: « Existe un ente tal que sólo él es ahora el rey de Francia, y este ente es calvo» , lo cual es falso. Aquí la negación

no es que el rey de Francia no es calvo (lo cual es falso), sino que «*No* existe ningún ente que sea el único que hoy es el rey de Francia, y que ese ente sea calvo». Esta proposición es cierta, de modo que se preserva la ley del tercero excluido.

La idea en síntesis: el lenguaje y la lógica

«Supongamos que cada cual tuviera una caja y que dentro hubiera algo que llamamos "escarabajo". Nadie puede mirar en la caja de otro, pero cada cual dice que sabe lo que es un escarabajo a partir del suyo. Sería perfectamente posible que cada uno tuviera algo distinto en su caja. Incluso podríamos imaginar algo que cambia constantemente. Pero ¿acaso diríamos que la palabra "escarabajo" tiene un uso en el lenguaje de estas personas? Si lo tuviera no sería el de designar una cosa. Lo que hay en la caja no tiene lugar en el juego de lenguaje; ni siquiera como *algo*: pues la caja podría incluso estar vacía … se neutraliza, sea lo que sea.»

¿Qué quieres decir cuando usas la palabra « dolor» ? Es obvio, piensas: te refieres a una sensación particular, a una más de las muchas otras cosas propias de tu experiencia subjetiva. Pero el filósofo austríaco Ludwig Wittgenstein sostenía que no es eso —no *puede ser* eso— lo que estás haciendo. Intentó explicarlo por medio de la analogía del escarabajo en la caja. Piensa en tu experiencia íntima como si fuera una caja; lo que sea que hay dentro lo llamas « escarabajo» . Todo el mundo tiene una caja, pero a lo sumo pueden mirar en la suya, nunca en la de nadie más. Todo el mundo usará la palabra « escarabajo» cuando hable de los contenidos de su caja, e incluso es perfectamente posible que las distintas cajas contengan cosas distintas o incluso nada en absoluto. Con « escarabajo» la gente simplemente dice « lo que sea que haya en la caja» : los contenidos reales son irrelevantes y no tienen nada que ver con el significado; el escarabajo mismo, sea lo que fuere, « queda fuera de consideración» . Cuando hablamos de lo que ocurre en nuestro interior usamos un lenguaje aprendido del discurso *público*, gobernado por reglas *públicas*. Las sensaciones íntimas, privadas, fuera del alcance del escrutinio de los otros, no pueden jugar ningún papel en esta actividad esencialmente pública; sean lo que sean, de hecho, estas sensaciones no tienen nada que ver con el *significado* de palabras como « dolor» .

Cronología

c. 375 a. C	*c.* 350 a. C.
¿Qué es el arte?	Formas de razonamiento

«Mira la frase como un instrumento, y su sentido como su uso.»

Ludwig Wittgenstein, 1953

El argumento del lenguaje privado Wittgenstein propuso la analogía del escarabajo en la caja al final de uno de los argumentos filosóficos más influyentes del siglo XX: el llamado argumento del « lenguaje privado» . Antes de Wittgenstein la concepción común (y del sentido común) del lenguaje era que el sentido de las

palabras proviene de representar o apoyarse en objetos del mundo; las palabras eran «denotativas» (eran nombres o etiquetas que designaban cosas asociándose a ellas). En el caso de sensaciones como el dolor —proseguía la teoría—, el proceso de denominación tiene lugar gracias a alguna forma de introspección, en la que un hecho mental particular, o una experiencia, se identifica y asocia con una palabra determinada. Además, para filósofos como Descartes y Locke, que adherían la «vía de las ideas» (véase página 16), según la cual *todo* nuestro contacto con el mundo está mediado por representaciones interiores o «ideas», el significado de todo el lenguaje debía ser en última instancia dependiente de un proceso interior, en el que cada palabra se une a uno u otro objeto mental. El argumento del lenguaje privado niega que las palabras puedan significar de este modo.

Supongamos (Wittgenstein nos invita a imaginar) que decidimos registrar cada una de las sensaciones particulares escribiendo en un diario la letra S, en la que S es un signo completamente privado que significa «la sensación que estoy teniendo ahora». ¿Cómo podrías saber en sucesivas ocasiones si habías aplicado el signo correctamente? La única cosa que hacía que tu designación fuera correcta la primera vez fue tu decisión de que podía serlo: pero lo único que la hace correcta en sucesivas ocasiones es la decisión que tomas en ese momento. Dicho de otro modo, puedes decidir lo que quieras: si la designación *parece* correcta, lo es; y «esto sólo significa que en este caso no podemos hablar de "correcto"». No existe un criterio independiente de «corrección», concluye Wittgenstein, no hay nada fuera de nuestra propia y subjetiva experiencia que pueda actuar como estándar; es como si alguien protestara diciendo «¡Pero yo sé cómo soy de alto!», mientras pone la mano encima de su cabeza para probarlo. Puesto que no existe un modo que no sea arbitrario para dar cuenta de si un signo privado se ha aplicado correctamente o no, un signo tal no puede tener significado; y un lenguaje hecho de estos signos (un «lenguaje privado») carecería de significado, sería ininteligible incluso para su hablante.

1690	1905	1953
El velo de la percepción	El rey de Francia es calvo	El escarabajo en la caja

Ayudar a la mosca a salir de la botella

Las repercusiones del argumento del lenguaje privado de Wittgenstein fueron mucho más allá de la filosofía del lenguaje. En la primera mitad del siglo XX, el lenguaje fue el centro de interés de la mayor parte de trabajos filosóficos, pues se asumió de modo generalizado que los límites del conocimiento estaban circunscritos en el lenguaje: «De lo que no se

puede hablar hay que callar», tal como dijo el joven Wittgenstein. De modo que este importante giro en la comprensión del lenguaje representó una sacudida violenta en la filosofía en su conjunto. Asimismo, el impacto de la obra de Wittgenstein en el estilo y el método filosóficos fue igualmente significativo.

Wittgenstein tenía la sensación de que la mayor parte de la filosofía moderna estaba esencialmente mal concebida, basada en una mala comprensión del lenguaje fundamental: la errónea manera de pensar que expuso a través del argumento del lenguaje privado. Pensaba que los filósofos otorgan demasiada importancia a formas particulares de expresión y, en cambio, no prestan la debida atención al uso del lenguaje en la interacción social real. Se han acostumbrado a abstraer y a generalizar para aislar la observación de determinados problemas, que intentan resolver; efectivamente se montan sus problemas por su propia cuenta, pues «el lenguaje se ha ido de vacaciones». Es célebre la advertencia de Wittgenstein: se trata de encontrar una terapia (a través de la filosofía), no una teoría. Según la pintoresca imagen de Wittgenstein, los filósofos son como moscas atrapadas en una botella: y su trabajo consistía en «mostrar a la mosca la salida de la botella».

«Si un león pudiera hablar, no lo entenderíamos.»

Ludwig Wittgenstein, 1953

El significado en el uso De modo que el significado de las palabras no funciona, no puede funcionar, del modo que supone el modelo del «proceso interior». ¿Cómo significan, pues? Comprensiblemente, tras demostrar la imposibilidad del lenguaje privado, Wittgenstein insiste en la necesidad del lenguaje *público*: las palabras sólo tienen significado «inscritas en la vida». Lejos de consistir en un misterioso proceso oculto que se produce en nuestro interior, el significado del lenguaje descansa en la superficie, en los pormenores del uso que hacemos de él.

El error consiste en suponer que deberíamos descubrir el uso y el propósito del lenguaje, y entonces investigar más a fondo para desenterrar (como una especie de hecho adicional) el significado. Porque el significado es algo que se establece *entre* los hablantes de un lenguaje: el acuerdo sobre el sentido de una palabra es esencialmente un acuerdo sobre su uso. El lenguaje es público, perfectamente entretejido en la fábrica de las vidas que vivimos juntos; compartir un lenguaje es compartir una cultura de creencias y supuestos, y compartir una perspectiva común del mundo.

Para elaborar la idea del significado como uso, Wittgenstein propone la noción de

«juego de lenguaje». El dominio del lenguaje reside en ser capaz de hacer un uso apto y diestro de las palabras y expresiones en distintos contextos, desde los ámbitos técnicos y profesionales muy precisos, hasta los amplios ruedos sociales. Cada uno de estos distintos contextos, general o especializado, constituye un juego de lenguaje distinto, en el que se aplica una serie de reglas específicas; estas reglas no son correctas o erróneas, pero pueden ser más o menos apropiadas a una función o a unos propósitos particulares en la vida.

La idea en síntesis: los juegos de lenguaje

33 Ciencia y pseudociencia

Los fósiles son los vestigios o las huellas de criaturas que vivieron en el pasado, que se fosilizaron tras su muerte y se han preservado en las rocas. Se han descubierto cientos de tipos distintos de fósiles…

1. …desde bacterias primitivas que vivieron y murieron hace tres billones y medio de años, hasta los primeros humanos, que aparecieron en África en los últimos doscientos mil años. Los fósiles, y su disposición en las sucesivas capas geológicas, son un tesoro oculto de información sobre el desarrollo de la vida en la Tierra que nos muestran cómo se desarrollan las últimas formas de vida a partir de las primeras.

2. …desde bacterias simples hasta los primeros humanos. Todas estas criaturas extinguidas, junto con todas las criaturas que viven hoy, fueron creadas por Dios en un plazo de seis días hace unos 6 000 años. La mayor parte de los animales fosilizados murieron en un diluvio universal catastrófico que tuvo lugar unos 1 000 años más tarde.

Dos concepciones del todo opuestas sobre cómo se originaron los fósiles y qué nos revelan. La primera es la concepción ortodoxa que nos darían la mayoría de los geólogos o de los paleontólogos. La segunda podría ofrecérnosla un creacionista, para el que el relato bíblico de la creación del universo que encontramos en el Génesis es literalmente verdadero. Al creacionista no le inspira ninguna simpatía el otro modo de ver las cosas: cree que los científicos ortodoxos están radicalmente equivocados en muchos aspectos cruciales, sobre todo en la aceptación de la teoría de la evolución por medio de la selección natural; el científico ortodoxo piensa que al creacionista lo guía su fervor religioso, tal vez políticamente motivado, y considera que está del todo engañado si pretende formar parte de una empresa científica seria. Pues el creacionismo, de acuerdo con el punto de vista científico mayoritario, es una superchería disfrazada de ciencia, es decir, « pseudociencia» .

Cronología

c. 350 a. C.	*c.* 1300
Formas de razonamiento	La navaja de Ockham

Si estás en un agujero…

La secuencia cronológica de la evolución requiere que nunca haya habido ninguna « alteración» geológica (que los fósiles no hayan ido a parar a un estrato equivocado). Se trata de una hipótesis completamente

comprobable y perfectamente falsable: bastaría con encontrar un solo fósil de dinosaurio en la misma piedra en que se halla un fósil de humano o de alguna herramienta para que la teoría de la evolución quedara en agua de borrajas. De hecho, entre los millones de fósiles que se han descubierto no se ha encontrado una sola alteración de las capas: una contundente confirmación de la teoría. Para el creacionista esta evidencia supone una complicación considerable. Entre las diversas tentativas desesperadas de encontrar una explicación convincente para la evidencia, una de las propuestas sugeridas es «una intervención clasificatoria hidráulica» por la que se supone que las diferentes densidades, formas, tamaños, etc., de los cuerpos dan lugar a distintos grados de profundidad y así separan a los animales en distintas capas. Otra idea es la de que los animales de menor tamaño estuvieron en mejores condiciones de escapar a tierras más elevadas y con ello evitaron ahogarse por más tiempo. Si estás en un atolladero geológico…

Cuestiones científicas ¿Qué es exactamente la ciencia? Es evidente que necesitamos una respuesta a esta pregunta si queremos reconocer a los impostores, y diferenciarlos de los que se dedican a esta actividad. En cualquier caso la pregunta es importante: la ciencia tiene mucho que reivindicar y difícilmente podría exagerarse su importancia. La vida humana se ha transformado completamente en el espacio de unos pocos cientos de años: se han erradicado enfermedades devastadoras; viajes que llevaban semanas pueden hacerse en horas; los humanos han aterrizado en la Luna; se ha descubierto la estructura subatómica de la materia. Debemos estos logros, junto con una miríada de muchos otros, a la ciencia. El poder transformador de la ciencia es tan grande que la mera reivindicación de algo como «científico» a menudo se esgrime para disuadir los análisis o las valoraciones críticas. Pero no todos los desarrollos de la ciencia dominante están exentos de crítica, mientras que algunas de las reivindicaciones que se hacen desde los márgenes de la ciencia —o desde la pseudociencia— pueden ser cautivadoras, autocomplacientes o del todo peligrosas. De modo que la capacidad para discernir la diferencia es crucial.

1670	1739	1962
Fe y razón	Ciencia y pseudociencia	Cambios de paradigma

El método hipotético La concepción habitual consiste en que el «método científico» es hipotético: se inicia con los datos procedentes de la observación y otros medios, y avanza a partir de ahí hacia la teoría, intentando brindar hipótesis que expliquen los datos en cuestión. Una hipótesis exitosa es la que soporta los

posteriores exámenes y genera predicciones que no podrían haber sido anticipadas de otro modo. El movimiento va así desde la observación empírica hasta la generalización, y si ésta es buena y sobrevive a exámenes sucesivos, puede ser aceptada finalmente como « ley de la naturaleza» universal, que cabe esperar que se cumpla en circunstancias similares, con independencia del momento y del lugar. Su inconveniente, que reconoció hace unos 250 años David Hume, es el llamado « problema de la inducción» (véase la página 115).

La falsación

Una respuesta importante al problema de la inducción fue la del filósofo de origen austriaco Karl Popper. Esencialmente, dicho filósofo aceptaba que el problema no podía resolverse, pero escogió esquivarlo. Así, sugirió que no existe ninguna teoría que pueda considerarse jamás probada, con independencia de las evidencias que haya para sostenerla; lo que ocurre más bien es que aceptamos una teoría hasta que es falsada (o cuestionada). De modo que, mientras que un millón de observaciones de ovejas blancas no puede confirmar la hipótesis general de que todas las ovejas son blancas, la aparición de una sola oveja negra basta para falsarla.

La falsabilidad también era para Popper criterio para distinguir la verdadera ciencia de sus imitaciones. Una teoría científica plenamente satisfactoria asume riesgos, aventura predicciones que pueden ser examinadas y desveladas como erróneas; por el contrario, la pseudociencia siempre se salvaguarda y presenta las cosas de un modo vago, confiando así no ser puesta en evidencia. El falsacionismo sigue siendo muy influyente, aunque muchos no acepten el que excluya la inducción del método científico, o el que asuma una relación un tanto simplista entre las teorías científicas y las evidencias (presuntamente neutrales u objetivas) en las que se basan.

La determinación de la teoría depende de la evidencia Otro modo de llegar a un punto muy parecido consiste en decir que la teoría científica siempre « depende» de las evidencias disponibles: la evidencia por sí sola nunca basta para escoger definitivamente una teoría en vez de otra. Es más, en principio siempre existe un número de teorías alternativas que permite explicar o « hacer encajar» una determinada serie de datos. La cuestión es, entonces, si las diversas calificaciones y añadidos *ad hoc* requeridos para apuntalar una teoría son más de

las que pueden oponérsele. Este proceso de adecuación y depuración es una parte característica de la metodología científica, pero si el peso de las evidencias contra una teoría es muy considerable, no quedará más alternativa (racional) que rechazarla.

El problema del creacionismo es que existe un auténtico tsunami de evidencias contra él. Veamos sólo dos ejemplos:

• Los métodos radiométricos, y otros medios para determinar la antigüedad de un material, en los que se basa la geología, la antropología y las ciencias de la Tierra deben desecharse completamente para poder acomodar la cronología de la Nueva Tierra.

• La disposición estratificada de los fósiles en las piedras y la espectacular ausencia de alteraciones (fósiles que aparecieran en lugares donde no deberían) —evidencias contundentes de la evolución— obligan a los creacionistas a hacer contorsiones extravagantes.

El creacionismo también plantea toda una batería de problemas propios. Por ejemplo, haría falta una fuente de agua extraordinaria para inundar todo el planeta, y hasta la fecha las posibilidades sugeridas (el choque de un cometa glacial, una bóveda de vapor sobre la atmósfera, algún depósito subterráneo) no resultan ni siquiera remotamente plausibles. A menudo, se reprocha al creacionismo el no asumir ningún riesgo: no propone los planteamientos audaces y falsables característicos de la ciencia. Tal vez fuera más justo decir que se limita a hacer algunas propuestas fantasiosamente arriesgadas que no se basan en ninguna evidencia atendible.

La idea en síntesis: la evidencia falsa las hipótesis

«Si he visto un poco más lejos es porque me he subido a hombros de un Gigante.» El célebre comentario de Newton a su colega científico Robert Hooke expresa con elocuencia la idea popular sobre el progreso de la ciencia. Se supone que el progreso científico es un proceso acumulativo en el que cada generación de científicos edifica sobre los descubrimientos de sus predecesores: un avance en colaboración —gradual, metódico, ininterrumpido — hacia una mayor comprensión de las leyes naturales que gobiernan el universo.

Tal vez sea una representación popular y atractiva, pero bastante equivocada según el filósofo e historiador norteamericano Thomas S. Kuhn. En su influyente libro de 1962, *La estructura de las revoluciones científicas*, Kuhn ofreció una visión del desarrollo científico como algo más accidentado y discontinuo: una historia irregular e intermitente del progreso, salpicada de crisis revolucionarias conocidas como « cambios de paradigma» .

La ciencia normal y la revolucionaria Según Kuhn, en un período de supuesta « ciencia normal» una comunidad de científicos que trabaja en consenso opera en el interior de un marco o de una « visión del mundo» que él llama « paradigma» . Un paradigma es una vasta red flexible delimitada de ideas y supuestos compartidos: métodos y prácticas comunes, pautas implícitas sobre temas adecuados para la investigación y la experimentación, técnicas probadas y modelos de evidencia acordados, interpretaciones que pasan de una generación a otra sin ser apenas cuestionadas, y más cosas similares. A los científicos que trabajan en el seno de un paradigma no les preocupa explorar nuevos territorios o iluminar nuevos senderos; por el contrario, están principalmente entregados a resolver desajustes provocados por el esquema conceptual, despejando las anomalías a medida que se producen, y asegurando y ampliando de forma gradual las fronteras del territorio.

Cronología

c. 1300	1739
La navaja de Ockham	Ciencia y pseudociencia

Verdad científica y relativismo científico

Un rasgo central de la representación que brinda Kuhn del cambio científico es que se encuentra culturalmente inscrito en un conjunto completo de factores históricos y de otro tipo. Aunque el propio Kuhn

tenía mucho interés en distanciarse de una interpretación relativista de su trabajo, su explicación sobre cómo se desarrolla la ciencia pone en duda la noción misma de la verdad científica, así como la idea de que la finalidad de la ciencia es descubrir objetivamente hechos acerca de la verdadera forma de las cosas en el mundo. Pues ¿qué sentido tiene hablar de verdad objetiva si cada comunidad científica establece sus propios objetivos y criterios de evidencia y prueba; filtra todo a través de una red de supuestos y creencias previas; y toma sus propias decisiones sobre qué debe preguntarse y qué cuenta como una respuesta adecuada? La visión habitual consiste en que la verdad de una teoría científica es una cuestión de hasta qué punto se mantiene en pie y convive con las observaciones objetivas y neutrales sobre el mundo. Pero tal como han mostrado Kuhn y otros autores, no existen hechos «neutrales»: no existe una frontera nítida entre la teoría y los datos; toda observación está «cargada» de teoría, cubierta por una espesa capa de creencias y de teorías previas.

Los pasatiempos de Kelvin

Por su propia naturaleza, los cambios de paradigma pueden provocar algunos tropiezos. En 1900, en un asombroso arrebato de vanidad, el célebre físico inglés lord Kelvin declaró: «Ya no queda nada nuevo por descubrir en física. Todo lo que queda por hacer son mediciones cada vez más precisas». Sólo unos pocos años más tarde, las teorías de Einstein de la relatividad especial y general, y la nueva teoría cuántica, usurparon el trono ocupado durante unos dos siglos por la mecánica newtoniana.

Un período de ciencia normal puede prolongarse durante muchas generaciones, tal vez durante varios siglos, pero finalmente los esfuerzos de quienes integran la comunidad producen un volumen de problemas y anomalías que empiezan a socavar y a poner en cuestión el paradigma vigente. Esto produce una crisis que alienta a algunos a mirar más allá del marco establecido y a empezar a fraguar un nuevo paradigma, con lo cual se produce una migración de científicos —que puede llevar años o décadas— desde el viejo paradigma hacia el nuevo. El ejemplo que utiliza Kuhn para ilustrar esto es la traumática transición entre la cosmovisión ptolemaica en la que la Tierra era el centro y el sistema heliocéntrico de Copérnico. Otro cambio de paradigma abrupto fue la sustitución de la mecánica newtoniana por la física cuántica y por la mecánica de la relatividad en las primeras décadas del siglo XX.

Uso y abuso público

El término «cambio de paradigma» es uno de los pocos términos académicos o técnicos que se ha convertido en una expresión de dominio público. La noción de un cambio radical en el modo de pensar y mirar las cosas de la gente es tan sugerente y suena tan bien que se ha abierto camino en los más variados contextos. Así, la invención de pólvora marca un cambio de paradigma en la tecnología militar; la penicilina en la tecnología médica; los aviones con motor en la aviación; las raquetas de fibra de vidrio en el tenis; y así sucesivamente. Y de un modo aún menos serio la expresión se ha convertido en parte del socorrido arsenal de los manuales de marketing. Naturalmente es irónico que la obra de Kuhn representase un cambio de paradigma en el modo de entender el progreso la de la ciencia que tenían los filósofos.

La falta de unidad de la ciencia

Durante mucho tiempo se ha dado por descontado que la ciencia es una empresa esencialmente unificada. Parecía razonable hablar de un «método científico» (único, una serie de procedimientos y prácticas bien definidas que en principio podían aplicarse a muchas disciplinas diferentes; y servir para especular sobre las posibilidades de algún tipo de amplia unificación de las ciencias, en el que todas las leyes y los principios podrían de algún modo desembocar en una estructura exhaustiva, omniabarcadora e internamente consistente). La clave de una unión semejante es supuestamente una explicación de las ciencias del todo reductora, pues sugiere que todo quedará subsumido finalmente bajo la física. Sin embargo los trabajos recientes han brindado una descripción más completa de la inscripción social y cultural de las ciencias, y han puesto un mayor énfasis en la esencial falta de unidad de la ciencia. Y, con ello, hemos advertido que la búsqueda de un método científico único es probablemente una quimera.

Las acusadas discontinuidades y dislocaciones que

«**No dudo de que, por**

supone la explicación de Kuhn han hecho que sea un tanto polémica como tesis histórica, pero no obstante ha resultado muy influyente entre los filósofos de la ciencia. Algo que ha resultado de particular interés ha sido la afirmación de que los distintos paradigmas son « inconmensurables» : las diferencias básicas de sus lógicas inmanentes implican que los resultados de un paradigma son efectivamente incompatibles con, o indemostrables a partir de, otro paradigma. Así, por ejemplo, si bien podemos suponer que los « átomos» del filósofo griego Demócrito no pueden compararse con los fisionados por Ernest Rutherford, la inconmensurabilidad supone que los átomos de Rutherford también son diferentes de los descritos por la teoría moderna de la mecánica cuántica. Esta discontinuidad lógica en el seno de la gran arquitectura de la ciencia contradice frontalmente la visión que había prevalecido antes de la época de Kuhn. Hasta entonces se había aceptado que el edificio del conocimiento científico se erigía de un modo seguro y racional sobre los fundamentos establecidos por los anteriores científicos. Kuhn suprimió de un plumazo la idea del progreso conjunto hacia una única verdad científica, y la sustituyó por un paisaje de aspiraciones y métodos científicos determinados localmente, y a menudo en conflicto entre ellos.

ejemplo, la mecánica newtoniana mejoró gracias a la de Aristóteles, y que la de Einstein mejoró gracias a Newton, como instrumentos para resolver el rompecabezas. Pero soy incapaz de encontrar en su sucesión una dirección coherente en el desarrollo ontológico.»

Thomas Kuhn, 1962

La idea en síntesis: la ciencia, evolución y revolución

35 La navaja de Ockham

Los círculos de los cultivos son formas geométricas surcadas en campos de trigo, cebada, centeno… tales formaciones, a menudo muy vastas y de diseños bastante intrincados, se han encontrado por todo el mundo en una cantidad creciente desde 1970. Como los medios de comunicación se hicieron mucho eco, en un principio se produjo una febril especulación acerca de su origen.

Entre las teorías con más predicamento están las siguientes:

1. Los círculos marcaban las zonas de aterrizaje de las naves espaciales extraterrestres, o de los OVNIS, que habrían dejado formas reconocibles en el terreno.
2. Los círculos los crearon bromistas humanos, que acudían por las noches, provistos de cuerdas y otras herramientas, para crear esas figuras y así llamar la atención de los medios de comunicación y provocar las especulaciones.

> La navaja de Ockham recibe el nombre de William Ockham, filósofo Inglés del siglo XIV. La «navaja» se debe a la idea de rasurar cualquier supuesto innecesario de una teoría.

Las dos explicaciones parecen adecuarse a las evidencias disponibles, así que ¿cómo decidir cuál de las dos teorías deberíamos creer? En ausencia de cualquier otra información, ¿podemos realizar la elección entre una teoría y sus rivales? De acuerdo con un principio conocido como la navaja de Ockham, sí podemos: cuando se nos ofrecen dos o más hipótesis para explicar un fenómeno determinado, es razonable aceptar el más simple: el que incluye menos supuestos no probados. La teoría 1 supone que los OVNIS existen, un supuesto del que no existen evidencias claras. La teoría 2 no supone ninguna actividad paranomal; además sólo supone un tipo de comportamiento humano, la travesura, que ha sido muy común a lo largo de la historia. De modo que estamos racionalmente justificados —de forma provisional y sólo hasta que una nueva evidencia aparezca— para creer que los círculos de los cultivos son obra de unos bromistas humanos.

Cronología

c. 350 a. c.	c. 1300	1637
Las formas del razonamiento	La navaja de Ockham	El problema del cuerpo y la mente

Caballos, no cebras

Para los doctores siempre es tentador, especialmente para los jóvenes, diagnosticar un caso exótico y raro allí donde el tópico y la explicación mundana son más verosímiles. Para corregir esta tendencia, a los estudiantes de medicina norteamericanos se les advierte a veces: « cuando oigáis el ruido de los cascos, no esperéis que aparezca una cebra» : la mayoría de las veces el diagnóstico más obvio es el más correcto, Sin embargo, como ocurre en aplicaciones parecidas de la navaja de Ockham, la explicación más simple no es necesariamente la mejor, y ningún doctor que no hubiera reconocido nunca a un caballo sería un doctor de caballos. Es evidente que los doctores norteamericanos que trabajan en África tienen que invertir sus aforismos.

De hecho, en este caso la navaja de Ockham funciona a la perfección. Hoy se sabe que la teoría 2 es la correcta porque los bromistas involucrados lo han admitido todo. Pero ¿es siempre tan eficaz la navaja de Ockham?

Ambiciones y limitaciones Conocido en ocasiones como el principio de parsimonia, la navaja de Ockham es en esencia un precepto de no recurrir a una explicación más complicada en los casos en que existe una más simple. Si diversas explicaciones alternativas están disponibles debemos, en igualdad de condiciones, dar preferencia a la más simple.

A veces se critica la navaja de Ockham por no cumplir con lo que, de hecho, se proponía no hacer. Las teorías empíricas siempre dependen de la determinación de los datos en los que se basan (véase la página 138), de modo que siempre existen diversas explicaciones posibles para un determinado conjunto de evidencias. El principio no sostiene que la explicación simple sea la conecta, sino sencillamente que es más probable que sea cierta y que ello explica que sea preferible hasta que existan razones fundadas para adoptar una alternativa más elaborada. Se trata esencialmente de una regla general o de un precepto metodológico, especialmente válido (se supone) para orientar los propios esfuerzos en los primeros estadios de una investigación.

1739	1912	1962	1981
La ciencia y la pseudociencia	Otras mentes	Cambios de paradigma	El cerebro en una cubeta

El principio del KISS

En ámbitos como la ingeniería u otras disciplinas técnicas la navaja de Ockham cobra la forma un tanto indecorosa de « principio del KISS (beso)» . En el desarrollo de programas informáticos, por ejemplo, la complejidad y la multiplicación de especificaciones ejercen una atracción irresistible, algo que evidencia el apabullante lote de « pijadas» que se añaden de forma ingeniosa y que puntualmente ignoran el 95% de los usuarios. Lo esencial del principio cuya aplicación pretende evitar estos excesos se resume entonces como: « Keep It Simple, Stupid» (No lo compliques, tonto).

La navaja en acción Aunque por lo general no se reconoce explícitamente, la navaja de Ockham se esgrime a menudo en los debates científicos y filosóficos, incluidos algunos de los mencionados en este libro.

El problema del cerebro en una cubeta (véase la página 8) establece dos situaciones enfrentadas, y las dos son aparentemente compatibles con las evidencias de que disponemos: somos seres reales en un mundo real, o cerebros en cubetas. ¿Es racional creer que somos lo primero en vez de lo segundo? Sí, de acuerdo con la navaja de Ockham, porque la primera es más *simple*: un solo mundo real, en vez de un mundo virtual creado en una cubeta, más la computadora que alimenta la cubeta, los científicos perversos, y el resto. Pero también aquí, como en otros casos, el problema se ha desplazado, no se ha resuelto: pues ¿cómo podemos decir cuál es la situación más simple? Podríamos, por ejemplo, insistir en que el número de objetos físicos es lo que importa y que, por lo tanto, un mundo virtual es más simple que uno real.

En una misma línea, a veces se desechan los demás problemas de la mente (véase la página 48) —el problema de cómo saber si los otros tienen mente— esgrimiendo la navaja de Ockham: existen muchas otras explicaciones posibles, pero es racional pensar que los otros tienen mente como nosotros mismos porque atribuirles pensamiento consciente es la explicación más *simple* para explicar su comportamiento. No obstante, una vez más las preguntas acerca de qué se considera lo más simple embotan considerablemente el filo de la navaja.

El asno de Buridán

Se supone que un uso juicioso de la navaja de Ockham facilita la elección racional entre dos teorías confrontadas. El asno de Buridán —que al parecer debemos a un discípulo de Guillermo de Ockham, Jean Buridan— ilustra el peligro de racionalizar excesivamente las elecciones.

El asno en cuestión, al encontrarse frente a dos almiares, no consigue dar con una razón para preferir un montón de paja a otro y, puesto que es incapaz de resolver la cuestión, no hace nada y muere de hambre. El fallo de la desdichada bestia consiste en suponer que el que no haya ninguna razón para preferir una cosa a otra hace que escoger sea irracional, y no hacer nada sea racional. Naturalmente, de hecho, es racional hacer *algo*, incluso si ese algo no puede determinarse mediante una elección racional.

La navaja también se utiliza a menudo contra una serie de explicaciones dualistas, pues se supone que es más simple no incorporar otro nivel de realidad u otro patrón de explicación, etc. La complejidad innecesaria —proponer un mundo mental y otro físico separados, y tener que hacer luego penosos malabarismos para explicar cómo se relacionan— es una de las principales objeciones que se hace al dualismo cartesiano del cuerpo y la mente. La navaja permite cortar una de las dos realidades, pero naturalmente no indica cuál de las dos es desechable. Los fisicalistas actuales (quienes suponen que todo —incluido nosotros— es en última instancia susceptible de explicación física) constituyen la gran mayoría, pero siempre quedan quienes, como George Berkeley, adoptan el patrón idealista (véase la página 19).

¿Una navaja sin filo? La idea de la simplicidad puede interpretarse de distintos modos. ¿Se trata de rechazar la incorporación de entidades injustificadas o de hipótesis injustificadas? Hay muchos aspectos distintos: a la limitación del número y de la complejidad de las hipótesis al mínimo se alude en ocasiones como «elegancia»; y a la minimización del número y la complejidad de entes, como «parsimonia». Pero cada una de estas opciones puede contradecirse: la incorporación de un ente desconocido, como un planeta o una partícula subatómica, puede suponer una gran oportunidad para el andamiaje teórico como para descartarla. Pero si existe una incertidumbre semejante acerca del sentido de la navaja, ¿es razonable esperar que nos proporcione una orientación segura?

La idea en síntesis: no lo compliques

«Ya he visto y oído antes una considerable insolencia cockney; pero nunca pensé que escucharía a un gallito petulante pedir doscientas guineas por echarle a la gente un tarro de pintura en plena cara.» Lamentablemente célebre es esta condena del crítico Victoriano John Ruskin ante el fantasmagórico cuadro de James McNeil Whistler titulado *Nocturno en negro y oro* de 1875. La cantidad de libelos que se sucedieron dejó en una victoria nominal aparente al artista —se le compensaron los prejuicios con un solo cuarto de penique—, pero en realidad sacó mucho más: una plataforma desde la que defender los derechos de los artistas para expresarse, liberados de las constricciones de la crítica, y lanzar el grito de guerra del esteticismo: «el arte por el arte».

La completa incomprensión de la obra de Whistler por parte de Ruskin no es inusual. Cada época actualiza la batalla entre el artista y el crítico, en la que el último —que a menudo refleja el gusto conservador del público— se lamenta con horror y desdén de los supuestos excesos de una nueva y desafiante generación de artistas. En nuestra época, es frecuente ver a los críticos llevarse las manos a la cabeza ante la última atrocidad artística: un tiburón envasado, un lienzo empapado en orines, una cama sin hacer. Es un eterno conflicto imposible de resolver porque su origen es una discrepancia sobre una de las preguntas más fundamentales: ¿qué es el arte?

De la representación a la abstracción Las concepciones que Ruskin y Whistler tienen de las propiedades que una obra de arte debe poseer tienen mucho que ver entre sí. Dicho en jerga filosófica, discrepan en la naturaleza del valor estético, el análisis del cual constituye la cuestión central en el ámbito de la filosofía que es la estética.

Cronología

c. 375 a. C.	*c.* 350 a. C.
La caverna de Platón ¿Qué es el arte?	La ética de la virtud

El ojo del observador

La pregunta más elemental, y también más natural, de la estética es si la belleza (u otro valor estético) se encuentra realmente «en» (o si es inherente a) los objetos a los que se les atribuye. Los realistas, y los objetivistas, sostienen que la belleza es una propiedad real que debe poseer el objeto, y por ello es enteramente independiente de las ideas o las respuestas de cualquiera respecto a ella; el *David* de Miguel Ángel

sería bello incluso si no existiera ningún humano que lo juzgara así, incluso si todo el mundo pensara que es espantoso. El antirrealista, o subjetivista, cree que el valor estético está necesariamente vinculado a los juicios y las respuestas de los humanos. Paralelamente a lo que ocurre con la pregunta acerca de si los valores morales son objetivos o subjetivos (véase la página 56), la pura excepcionalidad de la belleza, su estar «fuera del mundo» y su autonomía con respecto a los observadores humanos, puede forzarnos a una posición antirrealista, es decir, a considerar que la belleza depende del ojo del observador. No obstante, nuestras intuiciones parecen apoyar firmemente la sensación de que hay *algo* más que hace al objeto bello, aparte del mero hecho de que a nosotros nos lo parezca,

La kantiana idea de la validez universal parece confirmar esta intuición: los juicios estéticos están efectivamente basados sólo en nuestras respuestas y sentimientos subjetivos; pero tales respuestas y sentimientos están tan arraigados en la naturaleza humana que son universalmente válidos: cabe esperar que cualquier humano cabalmente constituido los comparta.

En la concepción de los griegos, el arte es una representación o un espejo de la naturaleza. Para Platón, la realidad última reside en un mundo de Ideas o Formas perfectas e inmutables (inextricablemente unido a los conceptos de bondad y belleza) (véase la página 12). El filósofo griego consideraba las obras de arte como un reflejo o una pura imitación de estas Ideas, inferior y poco fidedigno como modelo de verdad; por eso echó a los poetas y a otros artistas de su república ideal. Aristóteles compartía la concepción del arte como representación, pero adoptaba una concepción más favorable a sus objetos, considerándolos como un modo de completar algo que en la naturaleza sólo estaba realizado parcialmente, y por lo tanto como una vía para iluminar la esencia universal de las cosas.

1739	1946	1953
La teoria del abucheo y del hurra	La falacia intencional	El escarabajo en la caja

La teoría institucional del arte

«Me preguntaron cosas como: "¿Esto es arte?". Y yo contesté: "A ver…, si no es arte ¿qué narices se supone que hace en una galería de arte y por

qué viene la gente a verlo?”.»

Este comentario de la artista inglesa Tracey Emin refleja la « teoría institucional del arte» , muy discutida desde la década de 1970. La teoría sostiene que una obra de arte funciona como tal únicamente en virtud de si miembros autorizados del mundo del arte (críticos, galeristas, artistas…) le han otorgado tal título. A pesar de ser influyente, la teoría institucional ha tenido que afrontar muchas dificultades, la menor de las cuales no es la de proporcionar una información considerablemente pobre. Queremos saber *por qué* se consideran valiosas las obras de arte. Los miembros del mundo del arte deben tener *razones* para hacer los juicios que hacen. Si no, ¿cuál es el interés que debemos a sus opiniones? Y si las tienen, debería informársenos mejor de cuáles son.

La idea del arte como representación, y su estrecho vínculo con la belleza, ejerce un claro dominio en la modernidad. Pero, como reacción, en el siglo XX algunos pensadores proponen una aproximación « formalista» al arte, para la cual las líneas, los colores y otras cualidades formales son consideradas primordiales, y cualquier otra consideración, incluso los aspectos representacionales, es despreciada o excluida. Así, la forma prevalece sobre el contenido, preparando el camino para el abstraccionismo, que llegó a jugar un papel hegemónico en el arte occidental. Otra alternativa muy influyente en la representación fue el expresionismo, que renunció a cualquier observación atenta y fidedigna del mundo externo a favor de la exageración y la distorsión, mediante el uso de llamativos colores artificiales para expresar los sentimientos íntimos del artista. Instintivas y conscientemente no naturalistas, tales expresiones de las emociones y la experiencia subjetivas del artista fueron consideradas la marca distintiva de las verdaderas obras de arte.

«Podemos observar una compleja red de similitudes que se solapan y se entrecruzan; unas veces son similitudes generales, y otras similitudes de detalle.»

Ludwig Wittgenstein, 1953

Un aire de familia Un tema imperecedero de la filosofía occidental desde Platón ha sido la consecución de definiciones. Los diálogos socráticos plantean la característica pregunta —qué es la justicia, qué es el conocimiento, qué es la belleza—, y a partir de ahí proceden a demostrar, mediante una serie de preguntas y respuestas, que los interlocutores, a pesar de su presunto conocimiento, no poseen un verdadero conocimiento de los conceptos involucrados. El supuesto tácito es que el verdadero conocimiento de algo depende de la capacidad para definirlo, y esto es lo que los interlocutores de Sócrates (el portavoz de Platón) son incapaces de

hacer. Pero esto nos confronta a una paradoja, pues quienes son incapaces de proporcionar una definición de un determinado concepto suelen ser capaces de reconocer *qué no es*, lo que sin duda exige que sepan, en alguna medida, *qué es*.

El concepto de arte nos confronta con un caso de este tipo. Parece que sabemos qué es, por dificultoso que nos resulte definir las condiciones necesarias y suficientes para que algo se considere obra de arte. En nuestra perplejidad, tal vez sea natural preguntar si la tarea de definición no está en sí misma mal planteada: tal vez sea una pérdida de tiempo intentar identificar algo que se niega obstinadamente a dejarse asir.

La noción de aire de familia, que Wittgenstein expone en sus póstumas *Investigaciones filosóficas*, ofrece una salida de este laberinto. Consideremos la palabra «juego». Todos nosotros tenemos una idea clara de qué son los juegos: podemos dar ejemplos, comparar distintos juegos, decidir acerca de algunos casos dudosos, etc. Pero los problemas surgen cuando intentamos ir al fondo del asunto y encontrar algún significado esencial o alguna definición que los abarque todos. Pues no existe un denominador común: hay muchas cosas que los juegos tienen en común, pero no existe un único rasgo que todos ellos compartan. En resumen, no hay un significado esencial o un trasfondo oculto: nuestra comprensión de la palabra consiste ni más ni menos en nuestra capacidad para usarla de forma pertinente en una amplia variedad de contextos.

Si suponemos que «arte», como «juego», es una palabra que reúne cosas con un aire de familia, la mayor parte de nuestras dificultades se disipan. Unas obras de arte tienen muchas cosas en común con otras obras de arte: pueden expresar las emociones íntimas del artista; pueden destilar la esencia de la naturaleza; pueden conmovemos, repelernos o chocarnos. Pero si intentamos señalar algún rasgo que todas ellas compartan, estaremos buscando en vano; cualquier tentativa de *definir* el arte —de aislar un término que sea esencialmente fluido y dinámico en su uso— constituye un error y nos aboca al fracaso.

La idea en síntesis: los valores estéticos

Son muchos los que consideran a Richard Wagner como uno de los mayores compositores que han existido. Su genio creativo está fuera de duda; la procesión constante de peregrinos a su «santuario» en Bayreuth atestigua su enorme talento y su perdurable fascinación. Pero también parece indiscutible que Wagner fue un hombre excepcionalmente desagradable: de una arrogancia asombrosa y tremendamente obsesivo, carecía de cualquier escrúpulo cuando se trataba de explotar a los demás, era desleal con sus amigos más íntimos… el catálogo de debilidades y vicios es interminable. Y sus ideas eran, si cabe, incluso más repulsivas que su personalidad: intolerante, racista, virulentamente antisemita; fue un entusiasta abogado de la limpieza racial que clamó la expulsión de los judíos de Alemania.

¿Cuán decisivos son estos aspectos? ¿Tiene alguna relevancia nuestro conocimiento del carácter de Wagner, de sus inclinaciones, de sus ideas, etc., para nuestra comprensión y apreciación de su música? Podríamos suponer que estas consideraciones son relevantes en la medida en que informan o afectan a su obra musical; que saber lo que le motivó a producir una obra particular o cuáles eran las intenciones que propiciaron su creación podría darnos una comprensión más completa de sus logros y de su significado. Sin embargo, de acuerdo con una influyente teoría desarrollada a mediados del siglo XX, la interpretación de una obra debería ceñirse a sus cualidades objetivas; debería desatenderse rigurosamente todo factor externo o extrínseco (biográfico, histórico, etc.) relacionado con el autor de la obra. El (presunto) error de suponer que el significado y el valor de una obra pueden determinarlo factores semejantes se denomina « falacia intencional» .

Cronología

c. 375 a. C.	*c.* 350 a. C.
¿Qué es el arte?	Virtudes éticas

Obras públicas Aunque la idea se ha incorporado después a otros ámbitos, la falacia intencional procede originalmente de la crítica literaria. El término lo usaron por primera vez en un ensayo de 1946 William Wimsatt y Monroe Beardsley, dos miembros de la escuela de la Nueva Crítica que surgió en Estados Unidos en 1930.

«No es necesario conocer las intenciones personales del artista. La obra nos las cuenta.»

Susan Sontag, **n. 1933**

La principal preocupación de los nuevos críticos era que los poemas y otros textos fueran considerados como autónomos y autosuficientes; su significado

debería determinarse únicamente sobre la base de las propias palabras (las intenciones del autor, explícitas o tácitas, eran irrelevantes para el proceso de interpretación). Una obra, una vez liberada del mundo, se convierte en un objeto público al que nadie, ni siquiera el autor, tiene un acceso privilegiado.

¿Puede ser bueno el arte inmoral?

Existe un antiguo debate filosófico que se centra en la pregunta sobre si un arte moralmente malo puede ser bueno (desde el punto de vista estaban artístico). La pregunta suelen suscitarla figuras como Leni Riefensthal, la cineasta alemana cuyos documentales *El triunfo de la voluntad* (sobre los mítines de Nuremberg) y *Olimpia* (sobre las Olimpiadas de Berlín en 1936) eran esencialmente propaganda nazi pero los cuales, no obstante, se consideran a menudo como brillantes desde el punto de vista técnico y artístico. Los griegos antiguos habrían desechado la pregunta rápidamente, pues para ellos las nociones de belleza y de bondad moral íntimamente unidas, pero para los modernos resulta más problemática. Los propios artistas suelen ser relativamente indulgentes, entre ellos Ezra Pound, cuya posición resulta bastante característica: « El buen arte, por "inmoral" que sea, sigue siendo completamente virtuoso. El arte bueno no puede ser inmoral. Por arte bueno entiendo un arte que da cuenta de la verdad» .

La atención a la falacia intencional no es un asunto sólo teórico: pretendía ser un correctivo a las tendencias hegemónicas de la crítica. En efecto, en la medida en que nos atañe a los lectores profanos, está claro que dependemos de todo tipo de factores extraños al interpretar un texto; simplemente parece inadmisible suponer que nuestra lectura de un libro sobre la trata de esclavos sería la misma con independencia de si el autor es africano o europeo. Otra cosa es, naturalmente, si ello debería tener algún efecto, pero tal ve: deberíamos tener cuidado con las ideas que nos obligan a actuar de un modo tan distinto a la práctica común. Además, es incluso dudoso que sea posible, y mucho menos deseable, separar completamente la mentalidad de un autor y sus obras. Comprender las acciones de una persona implica suponer algunas cosas acerca de sus intenciones; ¿acaso la interpretación de una obra no depende en parte de semejantes supuestos e inferencias? A fin de cuentas, resulta complicado tragarse la idea de que lo que

un autor o un artista pretenden expresar con su obra es *irrelevante* para entender lo que realmente significa.

Imitaciones, falsificaciones y restos

Los peligros de la falacia intencional nos advierten de la conveniencia de ignorar las intenciones del creador cuando juzgamos el valor y el significado de una obra de arte. Pero, por más que nos veamos obligados a mirar una obra de arte aisladamente, desvinculada de las intenciones de su creador, podemos intentar mantener algunas distinciones que lamentaríamos (o nos sorprendería al menos) perder.

Supongamos que un falsificador creara un Picasso perfecto: con el estilo exacto del maestro, cada una de cuyas pinceladas fuera impecable, imposible de ser identificado por los expertos como una imitación. Normalmente subestimamos la copia, por buena que sea, puesto que no es la obra de un maestro; es una servil imitación, desprovista de originalidad y de genio creativo. Pero en cuanto se separa la obra de su origen, ¿acaso tales consideraciones no son pura cháchara? Un cínico podría decir cháchara por no decir algo peor: preferir un original es una mezcla poco edificante entre el esnobismo, la codicia y el fetichismo. La falacia intencional es un antídoto contra todas estas cosas, un recordatorio del verdadero valor del arte.

¿Y qué ocurre si no existen consideraciones que ignorar porque no existe un creador? Supongamos que millones de restos marinos imbricados al azar terminaran transformando una pieza de madera en una hermosa escultura, cuyo color, textura y equilibrio fueran perfectos. Conservaríamos esa pieza como un tesoro, pero ¿acaso sería una obra de arte, o simplemente arte? Parece evidente que no se trata de un artefacto. ¿Qué es entonces? Y ¿qué valor tiene? El hecho de que sea el producto de la creatividad humana cambia el modo en que lo contemplamos. ¿Pero no es acaso un error, si los orígenes de la escultura son irrelevantes?

Por último, supongamos que el mayor artista de la actualidad escogiera y dispusiera cuidadosamente un cubo y una fregona en una prestigiosa galería. Entonces, llegaría el empleado de la limpieza y dejaría su cubo y su fregona, idénticas a las del artista, al lado de la « obra de arte» . El valor artístico, en este caso, reside precisamente en el proceso de elección y disposición. Nada más diferencia a los dos cubos y las dos fregonas. Pero si consideramos sólo el carácter objetivo de los cubos y

las fregonas, ¿existe alguna diferencia?

Estas ideas sugieren que necesitamos examinar de nuevo nuestras actitudes hacia el arte. El peligro de que el vestido nuevo del emperador nos deslumbre es real.

La falacia de los afectos Al apreciar una obra de arte o un texto —especialmente si es complejo, abstracto o nos desafía de algún otro modo—, esperamos que los distintos tipos de público respondan de diferentes modos y tengan diversas opiniones. Esperamos que cada intérprete haga su propia interpretación, y en un sentido cada una de estas interpretaciones impone un significado distinto a la obra. Aparentemente, el hecho de que los diversos significados no puedan haber sido previstos por el artista parece dar la razón a la idea de la falacia intencional. Sin embargo, en su resuelto interés por las palabras mismas, la Nueva Crítica no estaba menos preocupada por excluir las reacciones y las respuestas del lector en la evaluación de una obra literaria. Al error de confundir el impacto que una obra puede tener entre el público con su significado lo llamaron la « falacia de los afectos» . Dadas las incontables respuestas subjetivas distintas que varias personas pueden experimentar, parece de poca ayuda vincularlas con el significado de la obra. Pero, una vez más, ¿acaso nuestra valoración de las cualidades supuestamente objetivas de una obra no puede estar influenciada por su capacidad para provocar diversas respuestas entre el público?

«El poema no le pertenece al crítico ni al autor (está desvinculado del autor desde su origen, y circula por el mundo fuera del alcance de sus pretensiones o de su control). El poema pertenece a los lectores.»

William Wimsatt y Monroe Beardsley, 1946

La idea en síntesis: los significados del arte

«Mira el mundo que te rodea: contempla su totalidad y cada una de sus partes, verás que no es otra cosa que una gran máquina, subdividida en un número infinito de máquinas menores, a su vez susceptibles de subdivisiones que superan lo que los sentidos y las facultades humanas pueden percibir o explicar. Todas estas diversas máquinas, e incluso sus partes más diminutas, se adecúan unas a otras con una precisión que cautiva hasta la admiración a cualquier hombre que las haya contemplado alguna vez. La curiosa armonía entre los medios y los fines, en toda la naturaleza, recuerda exactamente, aunque las supera con creces, las creaciones del ingenio humano; de los diseños humanos, del pensamiento, de la sabiduría y la inteligencia…

» … Y puesto que, por consiguiente, los efectos se parecen, debemos concluir, de acuerdo con todas las reglas de la analogía, que también las causas se parecen; y que el Autor de la Naturaleza es de algún modo similar a la mente del hombre, aunque dotada de muchas más facultades, proporcionadas a la grandeza de la obra que ha ejecutado. Mediante este argumento a posteriori, y únicamente mediante este argumento, se prueba simultáneamente la existencia de una Divinidad, y su similitud con la mente y la inteligencia humanas.»

Este sucinto planteamiento del argumento del diseño para probar la existencia de Dios lo pone en boca de su abogado Cleantes el filósofo David Hume en sus *Diálogos sobre la Religión Natural* que se publicaron póstumamente en 1779. El propósito de Hume es plantear el argumento para refutarlo (y la mayoría de los autores coinciden en que hizo un trabajo de demolición muy efectivo). Se trata, sin embargo, de un testimonio de la gran resistencia del argumento, y del atractivo inmediato que ejerce, no sólo capaz de sobrevivir al embate de Hume, sino de ir adoptando distintas formas y reapareciendo hasta la actualidad. Aunque la influencia que ejerció el argumento tuvo su auge en el siglo XVIII, sus orígenes se remontan a la Antigüedad y, de hecho, nunca ha pasado de moda desde entonces.

Cronología

c. 375 a. C.	*c.* 300 a. C.
El argumento del diseño	El problema del mal

En qué consiste el argumento La perenne vitalidad del argumento del diseño se debe a la poderosa y extendida intuición de que la belleza, el orden, la complejidad y la finalidad aparente que

La finalidad del mundo

El argumento del diseño del mundo también se conoce como

exhibe el mundo que nos rodea no puede ser *simplemente* el producto de los azarosos y ciegos procesos naturales. Sentimos que debe haber algún agente con una destreza y una capacidad mental inconcebibles, para planificar y realizar todas las maravillas de la naturaleza, tan exquisitamente diseñada y modelada para cumplir sus distintos papeles. Fijémonos, por ejemplo, en el ojo humano: es de una elaboración tan intrincada, y está adecuado a su función de un modo tan asombroso, que tiene que haber sido diseñado expresamente para ella.

« argumento teleológico» . El calificativo deriva de la palabra griega *telos*, que significaba « finalidad» o « propósito» , porque la idea que subyace al argumento es que la finalidad que nosotros, aparentemente, detectamos en el funcionamiento del mundo es una evidencia de que tiene un agente intencionado y responsable.

El argumento suele iniciarse con una lista de ejemplos elocuentes de tan asombrosa, en apariencia, habilidad de la naturaleza, para proseguir luego mediante una analogía con los artefactos humanos que evidencia claramente la impronta de sus creadores. Así, del mismo modo que un reloj, por ejemplo, está habilidosamente diseñado y construido para una finalidad particular, y nos permite inferir la existencia de un relojero, también las innumerables señales de la aparente intención y propósito en el mundo natural nos permiten concluir que en este caso existe asimismo un diseñador de la obra: un arquitecto a la altura del trabajo de diseñar las maravillas del universo. Y el único diseñador con poderes a la altura de la tarea es Dios.

1078	*c.* 1260	1670
El argumento ontológico	El argumento cosmológico	Fe y razón

El relojero divino y el ciego

En su *Teología natural* de 1802 el teólogo William Paley expuso una de las versiones más famosas del argumento del diseño. Si te encuentras un reloj en un matorral, inevitablemente deducirás de su complejidad y de la precisión de su construcción que debe haber sido obra de un relojero; del mismo modo, cuando contemplamos los portentosos ingenios de la naturaleza, nos sentimos obligados a concluir que también éstos deben tener un creador: Dios. Aludiendo a la imagen de Paley, el biólogo inglés Richard Dawkins describe los procesos de la selección natural como los de un « relojero ciego» , precisamente porque modela a ciegas las

complejas estructuras de la naturaleza, sin obedecer a ningún plan, ni propósito, ni finalidad.

Grietas en el diseño A pesar de su perenne atractivo, Hume y otros autores han planteado objeciones muy serias contra el argumento del diseño. Enumeramos a continuación las que más mella han hecho:

• Un razonamiento por analogía funciona afirmando que dos cosas son suficientemente parecidas en determinados aspectos conocidos como para justificar el que se les supongan semejanzas recíprocas en aspectos desconocidos. Las similitudes entre la psicología y el comportamiento de los humanos y los chimpancés son suficientemente numerosas como para suponer (aunque no podamos tener garantías) que, como nosotros, ellos experimentan sensaciones como el dolor. La fuerza de la analogía depende del grado de similitudes relevantes que existan entre las cosas comparadas. Pero los aspectos similares entre los artefactos humanos (por ejemplo, las cámaras) y los objetos naturales (por ejemplo, los ojos de los mamíferos) son de hecho relativamente irrelevantes, de modo que cualquier conclusión a la que se llegue por analogía será consecuentemente sesgada.

• El argumento del diseño parece vulnerable a una regresión infinita. Si la maravillosa belleza y la organización del universo requiere un diseñador, ¿cuánto más maravilloso será este universo de maravillas cuanto más lo sea el arquitecto que hay detrás de todo? Si hace falta un diseñador, parece que también debería haber un überdiseñador, y luego un über-überdiseñador, y luego… Así que, mientras que la eliminación de la regresión al infinito es una de las claves del argumento cosmológico (véase la página 160), en el argumento del diseño la amenaza de la regresión al infinito resulta perfectamente viciosa.

• Lo más cautivador del argumento del diseño es que explica cómo tales maravillas de la naturaleza como el ojo humano pueden existir y funcionar tan bien. Pero, precisamente, esas maravillas y su adecuación a la finalidad son las que resultan explicables a partir de la teoría de la evolución de Darwin y de la selección natural, sin necesidad de la intervención sobrenatural de un diseñador inteligente. Todo parece indicar que al relojero divino le ha quitado su trabajo el relojero ciego.

El afinamiento cósmico

Algunas variantes modernas del argumento del diseño surgen del asombro ante algo tan improbable como el que todas las condiciones en

el universo fueran exactamente como debían ser para que la vida pudiera desarrollarse y florecer. Si alguna de las muchas variables, como la fuerza de gravedad y la explosión inicial que expandió el universo, hubieran sido levemente distintas, la vida no habría surgido. En resumen, parecen existir evidencias de un afinamiento cósmico, tan preciso que debemos suponer que fue obra de un afinador inmensamente poderoso, Pero hasta las cosas más improbables ocurren. Es igualmente increíble que ganes un premio gordo de la lotería, pero es posible; y si lo ganaras no supondrías que alguien hubiera manipulado el resultado a tu favor: lo atribuirías a una extraordinaria suerte. Puede ser igualmente improbable que la vida se desarrolle, pero ello sólo se debe a que nosotros estamos aquí para advertir cuán improbable es… y para sacar conclusiones erróneas sobre la improbabilidad de algo que ha ocurrido.

• Incluso concediendo que los argumentos del diseño estuvieran justificados, no está nada claro hasta qué punto se sostienen. Muchos de los « artefactos» de la naturaleza podrían sugerir un diseño en equipo, de modo que podría resultar que hiciera falta un equipo de dioses en vez de limitamos a uno. Casi cada objeto natural, por impresionante que sea en general, está lejos de ser perfecto en los detalles; ¿los diseños defectuosos son acaso delatores de un diseñador defectuoso (no omnipotente)? Por lo general, el mal en el mundo y en sus objetos pone en duda la moral de su creador. Y naturalmente, no existe ninguna razón de peso para suponer que el diseñador, por bueno que fuera el trabajo hecho, siga con vida.

La idea en síntesis: el relojero divino

39 El argumento cosmológico

Pregunta: ¿Por qué hay algo en vez de nada?
Respuesta: Dios.
Éste es el inicio y el final del argumento cosmológico, y entre uno y otro no hay mucho más: es uno de los argumentos clásicos de la existencia de Dios, además de uno de los que mayor influjo ha tenido, y (según algunos) es asimismo el más dudoso de los argumentos de la historia de la filosofía.

De hecho, el « argumento cosmológico» es una clase o familia de argumentos, más que uno solo, pero todas las variantes son comparables por su forma, y tienen motivaciones similares. Todas están fundamentadas empíricamente, basadas (en su versión más familiar) en la aparentemente inobjetable observación de que todo lo que existe tiene alguna causa (véase el cuadro de la página siguiente). Esta causa es asimismo causada por otra causa, y así sucesivamente. Para evitar una regresión al infinito, hay que dar con una causa que no sea ella misma causada por ninguna otra causa: la causa primera, causa de sí y de todo, Dios.

¿Por qué existe algo? Si dejamos momentáneamente al margen la consideración de sus méritos, debe admitirse que el argumento cosmológico es una respuesta a la pregunta tal vez más natural, fundamental y profunda que nos cabe plantearnos: ¿por qué existe algo en vez de nada? Podría no haber existido nada, pero existe algo. ¿Por qué? Igual que el otro argumento clásico de la existencia de Dios, el argumento cosmológico tiene su origen en la Antigüedad, y constituye la base para las tres primeras vías del *Quinque Viae* (*Cinco Vías*), una serie de cinco argumentos para demostrar la existencia de Dios. Un cosmólogo moderno que preguntara « ¿Por qué existe algo?» sin duda nos remitiría al Big Bang, la gran explosión que se produjo hace algo así como 13 000 millones de años, y que dio origen al universo —a la energía, a la materia e incluso al tiempo —, lo cual sólo nos obligaría a reformular la pregunta: ¿qué, o quién, causó el Big Bang?

Cronología

c. 375 a. C.	c. 350 a. C.	1078
El argumento del diseño	Las formas del razonamiento	El argumento ontológico

Intentar no meter la pata La atracción del argumento cosmológico es que aborda una muy buena pregunta. Al menos, algo que parece una pregunta muy buena, y sin duda muy natural: ¿por

«En vez de proporcionarnos un argumento para la primera causa, nuestra experiencia

qué existimos (y por qué existe el resto del universo)? Pero ¿acaso el argumento cosmológico ofrece una buena respuesta? Hay unas cuantas razones para dudarlo.

es reticente a él.»

J. S. Mill, 1870

• La premisa aparentemente plausible en la que se basa el argumento cosmológico —que todo debe tener una causa— se explica por nuestra experiencia sobre cómo son las cosas en el mundo (o en el universo). Pero el argumento nos exige extrapolar esta idea a algo que se encuentra, por definición, *fuera* de nuestra experiencia, porque se encuentra fuera del universo; a saber, fuera de lo que sea que dio existencia al universo. En efecto, nuestra experiencia no puede arrojar luz ahí, y no está nada claro que el concepto sea ni siquiera coherente: el universo *significa* todo lo que existe, y su comienzo, si lo hubo, marca el comienzo del tiempo.

Variantes cosmológicas

La principal diferencia entre las distintas versiones del argumento cosmológico depende de un tipo particular de relación entre las cosas en las que se centra el interés. La versión más familiar, conocida también como el argumento de la causa primera, se atiene a una relación causal («todo es causado por alguna causa»), pero la relación puede ser de dependencia, o contingente, o explicativa o intelectiva. Se arguye que la secuencia de estas relaciones no puede extenderse indefinidamente, y para que la secuencia pueda hallar un fin, el punto de origen (es decir, Dios) debe carecer de las propiedades en cuestión. De manera que, según el argumento, Dios no debe ser causado (es causa de sí mismo); debe ser independiente de todo; no contingente (por tanto, existe necesariamente: no es posible que no hubiera existido); evidente; e inteligible sin referencia a nada más. (Por razones de simplicidad, en este artículo el argumento se plantea sólo en términos de la relación causal.)

c. 1260	1670	1739
El argumento cosmológico	Razón y fe	Ciencia y pseudociencia

El Dios de las lagunas

Históricamente se ha invocado al dios o a los dioses para explicar fenómenos de la naturaleza que superan la capacidad humana de

comprender y de conocimiento. Así, por ejemplo, en una época en la que se ignoraban las causas físicas de los fenómenos meteorológicos, como los truenos o los rayos, era habitual explicarlos como acciones divinas que expresaban enojo.

A medida que la ciencia ha avanzado y el conocimiento humano ha progresado, estas explicaciones han tendido a disminuir y a ser desplazadas. Antes de que Darwin propusiera la teoría de la evolución de las especies mediante la selección natural, el «dios de las lagunas» servía para explicar el aparentemente inexplicable orden y diseño del mundo natural (véase la página 156). En el caso del argumento cosmológico. Dios se retira hasta el extremo más remoto de la comprensión humana: al nacimiento del universo y al principio mismo del tiempo. En esta profunda fortaleza, Dios se encuentra fuera del alcance de las preguntas científicas. Pero ¿a qué precio? El reino de los cielos sí se ha contraído.

• Aparentemente, la principal premisa del argumento (todo tiene una causa) contradice la conclusión (existe algo, Dios, que no tiene una causa). Para evitar esta contradicción, Dios debe encontrarse fuera del ámbito de «todo», lo cual significaría entonces algo como «todas las cosas *de la naturaleza*». Dicho de otro modo, Dios debe ser sobrenatural. Esto puede satisfacer a quienes ya estaban persuadidos de la conclusión a la que supuestamente debía llevarnos el argumento. Para los demás —que son aquellos a los que debía convencerse— sólo aumenta el misterio y alimenta la sospecha de que los fundamentos del argumento son esencialmente incoherentes o incomprensibles.

• El argumento depende de la noción de que una regresión de las causas al infinito es intolerable: la cadena debe terminar en algún punto, y ese punto es Dios, que es causa de sí mismo. Pero ¿acaso la idea de una cadena infinita, y que implica que el universo no tiene un comienzo, es más complicada de digerir que un ser sobrenatural que se encuentra fuera del tiempo?

• Incluso si admitimos que la cadena de causas debe terminar en algún punto ¿por qué ese algo que no tiene otra causa que él mismo no podría ser el propio universo? Si la idea de la causa de sí mismo se acepta, Dios resulta redundante.

• El argumento cosmológico nos obliga a conferir a Dios un número de propiedades muy peculiares: debe carecer de causa (ser causa de sí mismo), necesariamente existente, y las demás. Pero son por sí mismas propiedades muy problemáticas y complicadas de interpretar. Lo que el argumento no *prueba* (suponiendo que pruebe algo) es que Dios posea las distintas propiedades

consistentes con la explicación teísta habitual: la omnipotencia, la omnisciencia, la bondad universal, etc. El Dios que surge del argumento cosmológico es muy extraño y tenue.

> **«Simplemente, el universo está ahí, y no hay más.»**
>
> **Bertrand Russell, 1964**

Entonces ¿qué *causó* el universo? El meollo del problema del argumento cosmológico es que, si la respuesta a la pregunta «¿Cuál es la causa del universo?» es X (Dios, por ejemplo, o el Big Bang), sigue siendo posible preguntar «Ya, pero ¿qué causó X?». Y si la respuesta a esta pregunta es Y, podemos seguir preguntando «¿Qué causó Y?». El único modo de impedir que la pregunta pueda retroceder hasta el infinito es insistir en que X (o Y o Z) es una causa de un tipo tal que la pregunta no puede plantearse. Y para ello es preciso que a X se le atribuyan algunas propiedades un poco misteriosas. A quienes son reticentes a aceptar esta consecuencia tal vez no les importe aceptar la implicación de que la cadena de causas se remonte al infinito, es decir, de que el universo no tenga un origen. O tal vez asuman el punto de vista de Bertrand Russell: el universo es, en última instancia, ininteligible, un hecho bruto del que no es posible hablar y sobre el que no cabe argumentar. Tal vez sea una respuesta insatisfactoria, pero no es peor que las otras disponibles para la mayoría de estas preocupaciones intratables.

La idea en síntesis: la causa primera y sin causa

Tómate un minuto para someter mentalmente a tus papilas gustativas el anacardo más grande imaginable: gordito, con su forma de «u» elegante y sinuosa, saladito y, lo mejor de todo, la textura: al masticarlo en tu boca se convierte en una pasta cremosa, blanda y suave. Todas las cualidades propias de un anacardo, cada una llevada al grado de su perfección. Ñam, ñam. ¿Lo ves en tu mente? Pues ahora viene lo mejor. Ese anacardo existe: el anacardo exquisito, que encama el grado más elevado de todas las perfecciones del fruto seco ¡existe de veras!

Porque lo que tenemos en mente es el mejor anacardo imaginable. Pero un fruto seco que exista en la realidad es, con toda seguridad, mejor que uno que sólo existe en nuestra mente. De modo que si el fruto seco en el que estamos pensando *sólo* existiera en nuestras cabezas, podría haber otro anacardo más rico, a saber, el que existe en nuestras mentes y en la realidad. Pero si éste existiera sucedería que sería posible imaginar un fruto seco más rico que el fruto seco más delicioso que quepa imaginar: y eso es una contradicción. De modo que el anacardo que imaginamos —el más rico que quepa imaginar— existe realmente: el anacardo insuperable debe existir, pues de lo contrario no sería insuperable.

De los frutos secos a Dios Lo que vale para los anacardos vale para Dios. O eso sugiere san Anselmo, el teólogo del siglo XI que acuñó la clásica formulación del argumento ontológico, uno de los argumentos a favor de la existencia de Dios más influyentes. Pasando completamente por alto los anacardos, san Anselmo empieza con la definición indiscutible (para él) de Dios como el ser « más excelso que cualquier cosa que pueda concebirse» . Pues bien, si podemos concebir perfectamente a Dios como tal, Dios debe existir como idea en nuestras mentes. Pero si Dios *sólo* existe en nuestras mentes, podríamos concebir a un ser todavía más excelso, a saber, uno que existiera en nuestras mentes y en la realidad. De modo que, a riesgo de incurrir en contradicción, Dios debe existir no sólo en nuestras mentes sino también en la realidad.

Cronología

c. 375 a. C.	*c.* 300 a. C.	1078
El argumento del diseño	El problema del mal	El argumento ontológico

Lógica modal y mundos posibles

La segunda afirmación del argumento ontológico de san Anselmo se parece mucho a la primera, pero la « existencia» se reemplaza ahora

por la « existencia necesaria» : la idea de que no puede concebirse que Dios *no exista*.

La existencia necesaria ha inspirado una serie de tentativas recientes (la más notable la de Alvin Plantinga) para reelaborar el argumento ontológico usando la lógica modal, en la que las ideas de posibilidad y necesidad se analizan en términos de mundos lógicamente posibles. Por ejemplo, supongamos que « absolutamente excelso» significase « existe y es omnipotente (etc.) en cualquier mundo posible» ; y concedamos que es al menos posible que un ser absolutamente excelso existiera (es decir, hay un mundo posible en el que semejante ser existe). Pero este ser que existe en un mundo posible *implica* que existe en todos los mundos, de modo que existe necesariamente. En lugar de aceptar esta conclusión, podemos preguntar qué concesiones hemos hecho para llegar a ella; en particular, que un ser absolutamente excelso deba existir *en cualquier* mundo posible. Pero negar esta posibilidad implica decir que un ser absolutamente excelso sea autocontradictorio. Así que ¿es posible que Dios, concebido como un ser absolutamente excelso, no tenga sentido?

A diferencia de la base empírica del argumento del diseño y del argumento cosmológico, el argumento ontológico se establece para probar, a priori y como si se tratara de un problema de necesidad lógica, que la existencia de Dios no puede negarse sin incurrir en contradicción: la idea de Dios implica su existencia. Exactamente igual que la comprensión del significado del concepto de un cuadrado implica cuatro lados, del mismo modo, arguye Anselmo, la comprensión del concepto de Dios implica saber que existe.

c. 1260	1670	1905
El argumento cosmológico	Razón y fe	El rey de Francia es calvo

¿La idea de Dios es incoherente?

Todas las versiones del argumento ontológico giran en torno a la idea de que podemos concebir un ser superior a cualquier otra cosa que podamos imaginar. Si no es de hecho posible (si el concepto de Dios resulta incoherente o ininteligible), todo el argumento se derrumba. Si el argumento debe probar la existencia de Dios tal como se ha concebido tradicionalmente (omnisciente, omnipotente, etc.), estas cualidades deben ser individualmente coherentes y compatibles entre sí, y cada una de ellas debe estar presente en Dios en el mayor grado posible.

No está nada claro que esto sea posible. Un dios omnipotente debería, por ejemplo, ser capaz de crear seres libres; un dios omnisciente excluye la posibilidad de que tales seres existan. Parece que la omnisciencia y la omnipotencia no pueden estar presentes al mismo tiempo en el mismo ser (un quebradero de cabeza considerable para la concepción tradicional de Dios). En el origen del problema del mal (véase la página 168) también se plantea el problema de si la idea tradicional de Dios es coherente.

Objeciones ontológicas Igual que el argumento cosmológico, el argumento ontológico constituye una familia de argumentos que comparten una única idea central. Todos son igualmente ambiciosos, pero ¿funcionan? La situación es complicada pues las distintas variantes del argumento afrontan diferentes tipos de crítica. Incluso san Anselmo presentó dos versiones distintas (en la misma obra). La versión que ya hemos dado (la primera formulación de san Anselmo del argumento y su planteamiento clásico) es vulnerable a dos líneas de ataque relacionadas entre sí.

Uno de los primeros críticos de san Anselmo fue un contemporáneo llamado Gaunilo, un monje de la abadía de Marmoutier en Francia. La preocupación de Gaunilo era que los argumentos como el ontológico podían utilizarse para probar que *nada* existe. Su propio ejemplo era una isla perfecta, pero el argumento funciona tan bien con los anacardos como con cosas inexistentes como las sirenas o los centauros. Evidentemente, si un argumento puede servir para probar la existencia de cosas inexistentes tiene un problema serio. Para hacer frente a esta línea de ataque, el defensor del argumento ontológico debe explicar por qué Dios es un caso especial (hasta qué punto difiere en aspectos relevantes de los anacardos). Algunos insisten en que las cualidades o «perfecciones» en las que reside la excelencia de Dios son literalmente perfectibles (susceptibles de alcanzar en principio un nivel más elevado) de un modo en que las propiedades del mejor anacardo no lo son. Si Dios es capaz de hacer todo lo que es concebible hacer, es omnipotente en un grado que no puede ser lógicamente superado; mientras que un anacardo gordito es un anacardo extraordinario, sigue siendo posible concebir uno más gordito y más extraordinario. De modo que la idea misma del anacardo más delicioso imaginable —a diferencia de Dios, el ser más poderoso que quepa imaginar— es incoherente. Así pues, el corolario, para que el argumento de san Anselmo funcione, es que el concepto de Dios debe formarse del todo de esas cualidades intrínsecamente perfectibles. Irónicamente, la aparente incompatibilidad entre estas cualidades tan similares amenaza con convertir el concepto de Dios en incoherente, minando todas las versiones del

argumento ontológico (véase el cuadro de la página anterior).

Gaunilo sólo tenía problemas con las artimañas verbales (creía que san Anselmo había conseguido definir la existencia de Dios). La misma preocupación parece subyacer al famoso ataque que Kant hace del argumento, en su *Crítica de la razón pura* de 1781. Su objeción es relativa a la consecuencia —explícita en la influyente reformulación cartesiana— de que la existencia sea una propiedad o predicado que puede adscribirse tanto a una cosa como a otra. La posición de Kant, completamente acorde con la lógica del siglo XX (véase la página 116), consiste en que decir que Dios existe no es atribuirle la propiedad de la existencia, junto con propiedades como la omnipotencia o la omnisciencia, sino afirmar que existe, de hecho, una instancia conceptual que reúne esas propiedades; y la verdad de tal predicado no puede ser determinada nunca a priori, sin observar cómo son, de hecho, las cosas en el mundo. En efecto, la existencia no es una propiedad sino una precondición de la posesión de propiedades. Tanto san Anselmo como Descartes cometen un error, la particularidad del cual se reconoce a la perfección si consideramos una afirmación como « Los anacardos que existen son más sabrosos que los que no existen» . San Anselmo realiza un salto ilícito desde un concepto hasta la instanciación del mismo concepto: primero asume que la existencia es una propiedad que algo puede tener o no tener; luego sostiene que tener esta propiedad es mejor que no tenerla; y por último concluye que Dios, puesto que es el ser más excelso que quepa imaginar, debe tenerla. Pero todo este prolijo edificio se desmorona de golpe si se niega a la existencia el estatus de predicado.

«Y definitivamente, que algo que es lo más excelso que cabe concebir, no puede existir sólo en la comprensión; así que puede concebirse como algo que existe en realidad, lo cual es más excelso.»

San Anselmo de Canterbury, 1078

La idea en síntesis: el ser más excelso imaginable

41 El problema del mal

La hambruna, el crimen, los terremotos, la enfermedad: millones de personas morirán en el futuro, jóvenes vidas se extinguirán absurdamente, muchos niños quedarán huérfanos e indefensos, la agonía de jóvenes y viejos será indiferente. Si estuviera en tu mano chascar los dedos y detener este catálogo de miserias, y no lo hicieras, serías un monstruo despiadado. Pero se supone que existe un ser que podría erradicarlo todo en un instante, un ser dotado de un poder, conocimiento y excelencia moral ilimitados: Dios. El mal está por todas partes, pero ¿cómo es posible que conviva con un dios que, por definición, tiene la capacidad de ponerle fin? Este asunto espinoso es el meollo del llamado «problema del mal».

El problema del mal es, sin duda, uno de los desafíos más serios que deben afrontar quienes quieren convencernos de la existencia de Dios. Frente a alguna calamidad terrible, la pregunta más natural es: « ¿Cómo puede Dios permitir que ocurra?» . La dificultad para dar con una respuesta puede poner a prueba seriamente la fe de los afligidos.

Se calcula que en 1984-1985 la sequía y la hambruna en Etiopía, agravadas por la inestabilidad política, provocaron aproximadamente un millón de agónicas muertes a causa del hambre.

¿Dios es ignorante, impotente, malvado o simplemente no existe? El problema surge como una consecuencia directa de las cualidades que se atribuyen a Dios en la tradición judeocristiana. Estas propiedades son esenciales a la concepción de Dios, y no es posible renunciar o modificar ninguna de ellas sin provocar grandes perjuicios en la concepción general. De acuerdo con el relato de la tradición teísta:

Cronología

c. 375 a. C.	c. 300 a. C.
El argumento del diseño	El problema del mal

¿Qué es el mal?

Aunque este asunto se suele conocer como « el problema del mal» , el término « mal» no es del todo adecuado. En este contexto la palabra refiere, de modo muy general, a todas las cosas malas que nos ocurren a todos y que son demasiado triviales como para ser calificadas de mal tal como se concibe normalmente. El dolor y el sufrimiento en cuestión se

deben tanto a causas humanas como naturales. Es habitual hablar del «mal moral» para englobar el sufrimiento causado por acciones inmorales de seres humanos (el asesinato, la mentira, y demás); y «mal natural» para englobar el sufrimiento causado por factores que escapan al control humano (desastres naturales como terremotos y enfermedades que no dependen de la actividad humana).

1. Dios es omnisciente: sabe todo lo que es lógicamente posible saber.
2. Dios es omnipotente: es capaz de hacer cualquier cosa que sea lógicamente posible hacer.
3. Dios es absolutamente bondadoso: está dotado de una buena voluntad universal y desea hacer cualquier cosa buena que sea posible hacer.

Con respecto al problema del mal, de estas tres propiedades básicas, pueden inferirse plausiblemente las tres siguientes ideas:

4. Si Dios es omnisciente, es perfectamente consciente de todo el sufrimiento y el dolor que tiene lugar.
5. Si Dios es omnipotente, es capaz de prevenir todo el dolor y el sufrimiento.
6. Si Dios es absolutamente bondadoso, debe desear prevenir todo el dolor y el sufrimiento.

En enero de 2007 Joshua DuRussel, un niño de 7 años de Michigan (EE. UU.), murió menos de un año después de que los médicos le descubrieran un tumor canceroso raro e inoperable que destruía progresivamente su tallo encefálico. Según un funcionario de su colegio, este niño que jugaba al básquet y era amante de los animales «tenía serias dificultades, pero nunca abandonó la esperanza y nunca se quejó».

El 8 de octubre de 2005 un terremoto catastrófico asoló la región Pakistani de Cachemira, y destruyó muchas ciudades y pueblos. Las cifras oficiales de muertos alcanzaron el número de 75 000; hubo otros 100 000 heridos y cerca de tres millones de persona perdieron sus hogares.

Si las proposiciones 4 y 6 son verdad, y Dios, de acuerdo con la definición de las proposiciones 1-3, existe, se sigue entonces que no habrá dolor ni sufrimiento en el mundo, porque Dios habrá seguido sus inclinaciones y lo habrá prevenido. Pero existe —manifiestamente— dolor y sufrimiento en el mundo, de modo que debemos concluir o bien que Dios no existe, o que

carece de una o más de las propiedades establecidas en las proposiciones 1-3. En suma, el problema del mal parece tener la implicación, especialmente indigesta para el teísta, o bien de que Dios no sabe lo que ocurre, o no le importa, o no puede hacer nada para evitarlo; o bien de que no existe.

c. 400 d. C.	1078	1670
La defensa del libre albedrío	El argumento ontológico	Razón y fe

Escurrir el bulto Las tentativas para evitar esta conclusión demoledora implican menoscabar alguno de los aspectos enumerados del argumento. Negar que exista en última instancia algo como el mal, tal como hacen los científicos cristianos, resuelve el problema de un plumazo, pero este remedio es demasiado difícil de tragar para la mayoría. Abandonar cualquiera de las tres propiedades básicas que se atribuyen a Dios (limitar su conocimiento, su poder o su excelencia moral) es demasiado oneroso para la mayoría de los teístas, de modo que la estrategia habitual es intentar explicar cómo pueden coexistir de hecho el mal y Dios (con todas sus propiedades intactas). Tales tentativas suelen implicar atacar la proposición 6 reivindicando que existen « razones morales suficientes» por las que Dios podría no siempre escoger eliminar el dolor y el sufrimiento. Lo que subyace a esta idea es la previa asunción de que, en alguna medida, Dios escogería esto en nuestro propio beneficio, a largo plazo. En suma, el advenimiento del mal en el mundo es, en última instancia, bueno: las cosas son mejores de lo que lo hubieran sido si no existiera el mal.

Dos problemas del mal

El problema del mal puede cobrar dos formas bastante distintas, aunque relacionadas entre sí. En la versión lógica (que consiste aproximadamente en lo explicado en la primera parte de este capítulo), la imposibilidad de que el mal y Dios coexistan se demuestra mediante un argumento deductivo: se afirma que el carácter de Dios es incoherente con el advenimiento del mal, y en consecuencia que la creencia en Dios es, de hecho, irracional. La versión del problema del mal que pone en evidencia esta incoherencia es, en efecto, una inversión del argumento del diseño (véase la página 156), pues usa la existencia de la interminable cantidad de horrores en el mundo para argumentar la improbabilidad de que sea la creación de un dios todopoderoso y lleno de amor, Esta segunda versión es menos ambiciosa que la versión lógica, y sólo pretende advertir que es improbable que Dios exista, aunque sea complicado rebatirlo. La versión lógica se desmonta formalmente

mostrando que la coexistencia de Dios y el mal es simplemente posible, por improbable que pueda parecer. La segunda versión presenta un gran desafío para el teísta, quien debe explicar cómo es posible que del elenco de males del mundo surja un bien mayor para los hombres.

En marzo de 2005 en Florida (Estados Unidos) se encontró enterrado en un pequeño foso el cuerpo medio descompuesto de Jessica Lunsford, una niña de nueve años. La había estrangulado, tras haberla secuestrado y violado varias semanas antes, John Couey, un hombre de 46 años que había cumplido pena por delitos sexuales.

Pero ¿exactamente qué bienes mayores se ganan al precio del dolor y el sufrimiento humanos? Es probable que la réplica más poderosa al problema del mal sea la llamada «defensa del libre albedrío», según la cual el sufrimiento en la Tierra es el precio que pagamos —y un precio perfectamente asumible— por nuestra libertad para elegir auténticamente nuestros actos (véase la página 172). Otra idea importante es la de que el verdadero carácter moral y la virtud se forjan con el sufrimiento humano: sólo sobreponiéndose a la adversidad, ayudando a los oprimidos, resistiendo a los tiranos, etc., puede brillar en todo su esplendor el valor real del héroe o del santo. Las tentativas de eludir el problema del mal tienden a topar con dificultades para explicar la arbitrariedad de la distribución y la magnitud del sufrimiento humano, pues, a menudo, son los inocentes los que más sufren mientras que los malvados salen indemnes; de modo que la cantidad de sufrimiento suele ser desproporcionada con lo que razonablemente requeriría la formación del carácter. Frente a la espantosa miseria, el último recurso del teísta suele consistir en alegar que «los designios del señor son inescrutables»: resulta insolente y presuntuoso que la débil mente humana ponga en duda los propósitos y las intenciones de un dios todopoderoso y omnisciente. Se trata, en efecto, de una apelación a la fe (es irracional invocar la razón para explicar las obras de la voluntad divina), y como tal es improbable que pueda consolar a quienes no están persuadidos.

La idea en síntesis: ¿por qué permite Dios que ocurran cosas malas?

42 La defensa del libre albedrío

La presencia del mal en el mundo representa el desafío más grande a la idea de que existe un dios omnipotente, omnisciente y lleno de amor. Pero los teístas dicen que el mal existe porque tomamos nuestras propias decisiones. El libre albedrío humano es un regalo divino de un inmenso valor, pero Dios no podría haberlo hecho sin asumir el riesgo de que lo usáramos mal. De modo que no puede ser responsable de las cosas malas que ocurren, pues solamente son culpa nuestra, y no deberían utilizarse para cuestionar la existencia de Dios.

La existencia manifiesta del mal —el drama cotidiano del dolor y del sufrimiento que nos rodea— sugiere que, si existe un dios, está muy lejos del ser perfecto que describe la tradición judeocristiana. En cambio, tiene más sentido suponer un ser que, o bien no está dispuesto, o bien es incapaz de evitar las cosas horribles que ocurren, y en consecuencia un dios que apenas merece nuestro respeto, y menos aún nuestra adoración.

Los intentos de cerrar el paso a este desafío requieren mostrar que, en efecto, existen suficientes razones por las que un dios perfecto en su dimensión moral podría haber escogido permitir la existencia del mal. Históricamente, la propuesta más popular e influyente es la llamada « defensa del libre albedrío» . Nuestra libertad para escoger de un modo genuinamente libre nos permite llevar vidas con un valor moral auténtico y entablar con Dios una relación del todo verdadera y amorosa. Pero podemos usar mal nuestra libertad y escoger mal. Era asumir un riesgo complicado y pagar un precio caro, pero Dios no podría haber eliminado la posibilidad de la bajeza moral sin privarnos de un gran regalo: la capacidad de la bondad moral. A pesar de la longevidad y de la perennidad de su atractivo, la defensa del libre albedrío plantea algunos problemas considerables.

Cronología

c. 300 a. C.	c. 400 a. C.	1078
El problema del mal	La defensa del libre albedrío	El argumento ontológico

El mal natural Tal vez la dificultad más obvia con que topa la defensa del libre albedrío es la existencia del mal natural en el mundo. Incluso si aceptamos que el libre albedrío es un bien precioso cuyo coste es el mal moral —el mal y las cosas malvadas surgen cuando los individuos usan su libertad para escoger mal—, ¿qué sentido puede tener el mal natural? ¿Cómo es posible que Dios socave o malogre nuestro libre albedrío de un modo u otro cuando de pronto nos barren el virus del SIDA, las hemorroides, los mosquitos, las inundaciones y los terremotos? La

seriedad de esta dificultad la ilustran algunas de las respuestas teístas: los desastres naturales, las enfermedades, las pestes y todas las cosas similares son, literalmente, obra del diablo, y de un montón de otros demonios y ángeles caídos; o tales aflicciones son un «justo» castigo de Dios al pecado original de Adán y Eva en el jardín del Edén. El último remedio remite todo el mal natural al primer episodio de mal moral, y con ello pretende exonerar a Dios de cualquier culpa. Pero esta explicación no resulta muy convincente. ¿No constituye una injusticia monstruosa que Dios mantenga el castigo sobre los tatara-(tatara-tatara-…) tataranietos de los pecadores originales? Y ¿cómo beneficia a quienes son juzgados de antemano por las acciones de sus (lejanos) antepasados el que se les otorgue el libre albedrío?

En la cultura popular

En la película del año 2002 *Minority Report*, Tom Cruise interpretaba al jefe de policía John Anderton del departamento de anticipación criminal de la ciudad de Washington. Anderton arrestaba a criminales antes de que cometieran el crimen, puesto que estaba convencido de que podían anticiparse sus actos con absoluta certeza. Pero cuando el propio Anderton es acusado, se convierte en un fugitivo, y le resulta imposible creer que sea capaz de asesinar. Al final, la posibilidad de anticipar los crímenes queda desacreditada, y con ello el determinismo, dejando intacta la fe de los espectadores en el libre albedrío.

1670	1789	1976
Razón y fe	Teorías del castigo	¿Es malo tener mala suerte?

¿Somos realmente libres?

El problema de la libertad implica que reconciliemos la concepción que tenemos de nosotros como sujetos libres con perfecto control sobre nuestros actos, con la comprensión (y de todo lo demás) que sugiere la ciencia. Dicho de forma simple, la idea del determinismo es que cualquier acontecimiento tiene una causa previa; cada estado del mundo requiere o está determinado por un estado previo que es, a su vez, el efecto de una secuencia de otros estados anteriores. Pero si todos nuestros actos están determinados de este modo por una serie de acontecimientos que se remontan atrás indefinidamente, a una época anterior incluso a nuestro nacimiento, ¿cómo es posible que nos

consideremos los verdaderos autores de esos actos y decisiones?

El determinismo parece amenazar la noción misma de actuar con libertad, y asimismo la noción de nuestro estatus de seres morales. Se trata de un problema profundamente decisivo que ha provocado una vasta variedad de respuestas filosóficas. Entre las muchas que existen, seleccionamos a continuación las principales:

• Deterministas radicales

Sostienen que el determinismo es cierto y que es incompatible con el libre albedrío. Nuestros actos están causalmente determinados y la idea de que son libres, en el sentido de que podríamos haber actuado de otro modo, es ilusoria. La censura o la alabanza morales, tal como se suelen entender, no son pertinentes.

• Deterministas moderados

Aceptan que el determinismo es cierto, pero niegan que sea incompatible con el libre albedrío, El hecho de que pudiéramos actuar determinista de esos mismos actos de un modo distinto en caso de tener la ocasión brinda una noción de la libertad de acción suficiente y satisfactoria. Es irrelevante que una decisión esté causalmente determinada; lo importante es que no sea forzada o contraria a nuestros deseos. En este sentido, una acción libre es susceptible de una valoración moral normal.

• Liberales

Están de acuerdo en que el determinismo es incompatible con el libre albedrío y, en consecuencia, rechazan el determinismo. El determinista moderado sostiene que es irrelevante que hubiéramos podido actuar de un modo distinto si hubiéramos elegido, porque una elección distinta también estaría causalmente determinada (o debería haberlo estado si el determinismo es cierto). Así, el liberal sostiene que el libre albedrío humano es real, y que nuestras decisiones y acciones no están determinadas. El problema para los liberales es explicar cómo es posible que tenga lugar una acción sin determinación (en particular, cómo es posible que un acontecimiento sin causa pueda no ser azaroso, pues el azar resultaría tan oneroso a la idea de la responsabilidad moral como el determinismo). Surge entonces la sospecha de una profunda laguna en el seno del liberalismo: tal vez el liberal ha desechado otras explicaciones sobre la acción humana sin poner en su lugar más que una gran caja negra.

¿La teoría cuántica viene al rescate?

A la mayoría de los filósofos les resulta complicado pasar por alto la idea del determinismo radical, de modo que unas veces han aceptado que el libre albedrío es ilusorio, y otras se han enfrentado a él valerosamente intentando integrarlo de algún modo. Al mismo tiempo, las tentativas de los liberales para explicar cómo pueden tener lugar los acontecimientos sin causa, o indeterminadamente, tienden a parecer *ad hoc* o simplemente descabelladas. Pero ¿la mecánica cuántica no puede resultar de ayuda al liberal? De acuerdo con la mecánica cuántica los acontecimientos a nivel subatómico son indeterminados (cuestión de pura suerte, «ocurren sin más»). ¿Acaso no proporciona esto una escapatoria del determinismo? En realidad, no. La esencia de la indeterminación de los quanta es el azar, de modo que la idea de que nuestras acciones y decisiones son en un nivel profundo azarosas no contribuye en absoluto a salvaguardar la noción de la responsabilidad moral.

Al margen de la dificultad del mal natural, la defensa del libre albedrío desemboca inevitablemente en otras dificultades filosóficas complicadas, relativas a la forma misma del problema del libre albedrío. La defensa asume que nuestra capacidad para elegir es genuinamente libre en el sentido pleno del término: cuando decidimos hacer algo nuestra decisión no está determinada o causada por ningún factor externo a nosotros; tenemos la posibilidad de obrar de otro modo. Esta llamada explicación «liberal» del libre albedrío concuerda perfectamente con nuestro cotidiano sentido de lo que ocurre cuando actuamos y decidimos, pero muchos filósofos consideran que es insostenible teniendo en cuenta el determinismo (véase el cuadro). Y naturalmente, si la explicación liberal que subyace a la defensa del libre albedrío es insostenible, lo que se hunde inmediatamente es la defensa misma.

La idea en síntesis: libertad para equivocarse

43 Razón y fe

A pesar de algunas heroicas tentativas recientes, la mayoría de los filósofos coincidirían en que no hay quien pueda resucitar los argumentos tradicionales para probar la existencia de Dios. Sin embargo, a la mayoría de los creyentes no les preocuparía demasiado está conclusión. Su creencia no depende de tales argumentos y, difícilmente, podría verse afectada por su refutación.

Para los creyentes, los requisitos normales del discurso racional son inadecuados a los asuntos religiosos. La especulación y el razonamiento abstractos de la filosofía no son lo que les ha llevado a creer en primera instancia, y tampoco podría convencerlos de renunciar a la creencia. Además, es arrogante, afirman, suponer que nuestros esfuerzos intelectuales podrían esclarecer o hacer comprensibles los designios divinos. Creer en Dios, no es, en ultima instancia, cuestión de razón sino de fe.

Tal vez la fe sea ciega, pero no se trata « simplemente de creer» . Quienes ponen la fe por encima de la razón —el llamado « fideísmo» — sostienen que la fe es un patrón alternativo de verdad y que, en el caso de la creencia religiosa, constituye el camino correcto. Un estado de convicción, que se alcanza al cabo mediante la acción de Dios en el alma, exige no obstante un acto deliberado y libre de la voluntad a favor de la fe; la fe requiere un salto en el vacío. Los filósofos, en cambio, pretenden hacer una valoración racional de los argumentos posibles a favor de la creencia religiosa, para cribar y sopesar las evidencias, y alcanzar una conclusión a partir de ellas. El fideísta y el filósofo parecen, pues, involucrados en proyectos diametralmente opuestos. Y, dada la aparente distancia, ¿existe alguna posibilidad de acuerdo o de encuentro entre ambos?

Cronología

c. 375 a. C.	c. 300 a. C.
La teoría del mandato divino	El problema del mal

Abraham e Isaac

El relato bíblico de Abraham e Isaac ilustra bien la insuperable distancia que separa la fe de la razón. Abraham es presentado como el ejemplo arquetípico y paradigmático de la fe religiosa por la indiscutible buena voluntad con que obedeció los mandatos divinos, hasta el punto de estar dispuesto a sacrificar a su hijo, Isaac. Sin embargo, si se extrapola su figura del contexto religioso y se la considera desde el punto de vista racional, el comportamiento de Abraham parece el de un individuo

trastornado. Cualquier alternativa hermenéutica de la situación era preferible y plausible a la que él escogió (¿acaso estoy loco/he oído mal/acaso Dios me está poniendo a prueba/no será el diablo haciéndose pasar por Dios/no podrían darme la orden por escrito?), de modo que su comportamiento es simple y llanamente incomprensible para la racionalidad del que no está dispuesto a creer.

El balance general de la fe En manos de un fideísta, el hecho de que la creencia religiosa no pueda defenderse de forma adecuada a partir de fundamentos racionales se toma una virtud. Si existiera una vía (completamente) racional no sería necesaria la fe, pero como la razón no consigue encontrar una justificación, la fe permite salvar esa brecha. El acto necesario de voluntad del creyente vincula el mérito moral a la adquisición de la fe; y, al menos, quienes comparten la devoción que no cuestiona a su objeto la reverencian como simple y honesta piedad. Algunos de los encantos de la fe son bastante obvios: la vida cobra un sentido evidente, las tribulaciones de la vida hallan alivio, los creyentes disponen del consuelo de saber que algo mejor les espera tras la muerte, etc. Es evidente que la creencia religiosa da respuesta a algunas necesidades y preocupaciones humanas fundamentales, primordiales, y mucha gente mejora visiblemente, incluso se transforma, al adoptar un modo de vida piadosa. Al mismo tiempo, los símbolos y bondades de la religión han suministrado una inspiración y una riqueza culturales prácticamente ilimitadas.

«Quien comienza amando más al cristianismo que a la verdad proseguirá queriendo más a su propia secta o Iglesia que al cristianismo, y terminará queriéndose a sí mismo más que a cualquier otra cosa.»

Samuel Taylor Coleridge, 1825

1670 Razón y fe
1739 Ciencia y pseudociencia

La apuesta de Pascal

Supongamos que sentimos que la evidencia de la existencia de Dios simplemente no es concluyente. ¿Qué hacer? Podemos creer en Dios o no. Si decidimos creer y acertamos (es decir. Dios existe), ganamos la dicha eterna; y si nos equivocamos, perdemos muy poco. Por otro lado, si decidimos no creer y acertamos (es decir, Dios no existe), no perdemos nada pero tampoco ganamos demasiado; pero si nos equivocamos, nuestra pérdida es colosal (en el mejor de los casos

perdemos la posibilidad de la dicha eterna, y en el peor de los casos sufrimos la condena eterna). Mucho que ganar y muy poco que perder: hay que ser idiota para no apostar por la existencia de Dios. Este ingenioso argumento para creer en Dios, conocido como la apuesta de Pascal, fue propuesto por Blaise Pascal en sus *Pensamientos* de 1670: y tal vez sea ingenioso, pero es fallido. El problema evidente es que el argumento requiere que *decidamos* qué creer, y ése no es precisamente el modo como funciona la creencia. Aunque hay algo peor, y es que lo primero que sentimos al hacer la apuesta es que no tenemos suficiente información sobre Dios para poder apostar; y además hacer la apuesta correcta depende de tener un conocimiento detallado de las ventajas y los inconvenientes de Dios. ¿Y qué ocurre si se da el caso de que a Dios no le molesta demasiado que lo adoren, pero le molestan considerablemente los tipejos calculadores que sólo apuestan si pueden esperar sacar algún beneficio para sí mismos?

«Dejadnos sopesar qué ganamos y qué perdemos apostando por la existencia de Dios. Dejadnos valorar estas dos posibilidades. Si ganas, lo ganas todo; si pierdes, no pierdes nada. Apuesta pues, sin dudar, por su existencia.»

Blaise Pascal, 1670

Muchas de las cosas que el fideísta apuntaría en el haber de la fe figuran como deberes del filósofo ateo. Pero entre los principios más preciados del liberalismo secular, acuñado de un modo memorable por J. S. Mill, se encuentra la libertad de expresión y de pensamiento, que convive difícilmente con el hábito del asentimiento acrítico que ensalza el creyente piadoso (véase el cuadro). La devoción acrítica que valora el fideísta puede parecer al no creyente una forma de credulidad y superstición. La perfecta aceptación de la autoridad puede conducir a los individuos a caer bajo la influencia de sectas y cultos inescrupulosos, y ello puede desembocar a veces en el fanatismo más atroz. Tener fe en los otros es admirable siempre que los otros en cuestión sean admirables. Cuando se acalla a la razón, cualquier forma de exceso puede apresurarse a ocupar su lugar; y hay que admitir que a veces el sentido religioso y la simpatía son arrojados por la ventana, y sustituidos por la intolerancia, la superstición, el sexismo y algunas otras cosas peores.

J. S. Mill y la libertad intelectual

En su *Sobre la libertad* de 1859, en una apasionada defensa de la libertad de pensamiento y de expresión, John Stuart Mill advertía de los peligros

de una cultura intelectualmente represiva, en la que se disuadiera de los cuestionamientos y la crítica de las ideas recibidas, y en la que «los intelectos más activos y críticos» tuvieran que temer limitarse a «especular libre y audazmente sobre los asuntos elevados». El desarrollo mental se encontraría constreñido y la razón intimidada, y el fundamento de la verdad se debilitaría: «tolerar las opiniones … como un prejuicio, una creencia independiente del (y a prueba de) argumento; ésa no es la manera en que debe defender la verdad un ser racional … La verdad, la verdadera, no es más que una superstición más, accidentalmente unida a palabras que enuncian una verdad».

«**Creo para comprender.**»

Agustín de Hipona, *c*. 400

De modo que ya tenemos el balance general, con el debe y el haber en cada lado, y a menudo los activos que hay en un lado parecen pasivos en el otro. En la medida en que los métodos de recuento usados son distintos, las cuentas mismas carecen de sentido, y ésa suele ser la inevitable impresión que nos deja el debate entre creyentes y no creyentes. Normalmente hablan de cosas distintas, no consiguen establecer ningún espacio común y lo único que consiguen es no moverse ni un milímetro de su posición inicial. Los ateos se complacen en demostrar que la fe es irracional; y el creyente considera esta presunta prueba irrelevante y bastante fuera de lugar. Al final, la fe es irracional o no racional; se enfrenta, orgullosa y desafiante, a la razón y, en un sentido, consiste precisamente en eso.

La idea en síntesis: el salto de la fe

44 Libertad positiva y negativa

La libertad es una de esas cosas sobre las que todo el mundo está de acuerdo. Es importante, es buena y constituye uno de los ideales políticos más importantes, tal vez *el* más importante. La libertad es también una de esas cosas en las que nadie se pone de acuerdo. ¿Cuánta deberíamos tener? ¿Es necesaria alguna limitación para que prospere? ¿Cómo puede evitarse que tu libertad de hacer algo entre en conflicto con mi libertad de hacer algo distinto?

La discusión sobre la libertad, que constituye por sí misma un problema peliagudo, se complica aún más por discrepancias fundamentales acerca de su verdadera naturaleza. Planea sobre ella la sospecha de que tal vez no sea realmente nada: no sólo ocurre que la palabra « libertad» tiene muchos matices, sino que además puede referirse a una cantidad de conceptos distintos, aunque relacionados. Así, que estamos en deuda con el influyente filósofo del siglo XX Isaiah Berlin por haber arrojado luz en este turbio panorama. En el centro de su reflexión sobre la libertad subyace una distinción entre la libertad positiva y la negativa.

Dos conceptos de libertad « George está sentado con un vaso de coñac frente a él. Nadie le apunta con una pistola en la cabeza para obligarle a beber. No existe coerción ni impedimento: no hay nada que lo fuerce a beber ni nada que lo disuada. Tiene libertad para actuar como le plazca. Pero George es alcohólico. Sabe que beber le hace daño, que incluso podría matarlo. Podría perder a sus amigos, a su familia, a sus hijos, su trabajo, su dignidad, el amor propio… pero no puede evitarlo. Tiende su mano temblorosa y lleva el vaso a sus labios.»

Cronología

c. 1260 — Actos y omisiones

1953 — El escarabajo en la caja

Aquí están en juego dos tipos de libertad muy distinta. Solemos pensar en la libertad como la ausencia de restricciones externas o de coerción: somos libres en la medida en que no existan obstáculos que nos impidan hacer lo que queremos. Esto es lo que Berlin denomina « libertad negativa» ; es negativa por cuanto se define por la ausencia de algo (cualquier forma de coacción o de interferencia externa). En este sentido, George el alcohólico es completamente libre. Pero George no se controla. Se siente impulsado a beber, incluso a sabiendas de que le conviene más no hacerlo. No tiene completo control sobre sí mismo y su destino no está enteramente en sus manos. En la medida en que se siente impelido a beber, no tiene elección y no es libre. Lo que le falta a George es lo que Berlin llama « libertad positiva» : positiva porque lo que la define es algo

que debe poseer el sujeto (autocontrol, autonomía, capacidad para actuar de acuerdo con lo que racionalmente consideramos más conveniente). En este sentido, es evidente que George no es libre.

Libertad negativa Somos libres, en el sentido negativo que señala Berlín, en la medida en que nadie interfiera en nuestra capacidad para actuar como nos place. Pero al ejercer nuestra libertad, es inevitable que topemos con la libertad del otro. Cuando canto libremente a toda voz en el baño, te niego la libertad de disfrutar de una tarde tranquila. Nadie puede disfrutar de una libertad ilimitada sin usurpar la libertad de los otros, de modo que en la vida en sociedad se requiere algún grado de compromiso.

1959 Libertad positiva y negativa
1971 El principio de la diferencia

La posición de los liberales clásicos se define por el llamado « principio del perjuicio» . Su planteamiento más célebre es el del filósofo Victoriano J. S. Mill en su *Sobre la libertad*, donde estipula que debería permitirse a los individuos actuar de cualquier modo que no cause perjuicio a los otros; sólo cuando se causa perjuicio existe una justificación para que la sociedad imponga restricciones. De algún modo, podemos definir un espacio de libertad privada que es sacrosanto e inmune a las interferencias externas y a la autoridad. En este espacio se permite a los individuos satisfacer sus preferencias y sus inclinaciones personales sin cortapisas; y en un sentido político los individuos tienen libertad para ejercer algunos derechos o libertades inviolables: de expresión, de asociación, de conciencia, etcétera.

Mientras la comprensión negativa de la libertad propugnada por los liberales suele ser la hegemónica, en los países occidentales, al menos, sigue habiendo muchos aspectos problemáticos. En particular, cabe preguntarse si la libertad de la que goza un individuo que no tiene la posibilidad ni los recursos para hacer lo que es « libre» de hacer, merece realmente tal nombre. Ésta es la sombra que proyecta el hecho de que todo ciudadano norteamericano sea libre para convertirse en presidente. Ciertamente no existen barreras legales o constitucionales, de modo que en esta medida todos los ciudadanos tienen libertad para hacerlo; pero de hecho, muchos ciudadanos están excluidos porque carecen de los recursos necesarios, es decir, del dinero, de la educación y del estatus social. En suma, carecen de la libertad *sustantiva* para ejercer los derechos que *formalmente* poseen. Pero para reparar estas deficiencias con vistas a transformar la mera libertad formal en una libertad real, sustantiva, el liberal estaría obligado a permitir formas de intervención estatal que parecen más adecuadas a la interpretación positiva de la libertad.

Libertad positiva Mientras que la libertad negativa es libertad con *respecto a* las interferencias externas, la libertad positiva suele caracterizarse como libertad para alcanzar determinados fines; como una forma de estimulación que permite al individuo realizar su potencial, alcanzar una visión particular de su realización, lograr un estado de autonomía personal y autodominio. En un sentido político más amplio, la libertad en este sentido positivo se entiende como la liberación de la presión cultural y social que, de otro modo, impedirían el progreso hacia la autorrealización.

«**Al sujeto (una persona o un grupo de personas) se le debería permitir hacer o ser lo que fuera capaz de hacer o de ser, sin que interfirieran otras personas.**»

Isaiah Berlin, 1959

Mientras que la libertad negativa es esencialmente interpersonal, y existe como relación entre los individuos, la libertad positiva, en cambio, es intrapersonal: es algo que se desarrolla y se cultiva en el interior del individuo. Del mismo modo que en el interior de George existe un conflicto entre su conciencia más racional y sus apetitos más bajos, generalmente el concepto positivo de la libertad supone una división del yo en partes elevadas y partes inferiores. La consecución de la libertad está determinada por el triunfo (moral, racional) del preferible yo elevado.

El abuso de la libertad

«¡Oh, libertad! ¡Cuántos crímenes se han cometido en tu nombre!», exclamó madame Roland antes de que la ejecutaran, en el año 1793. Pero las atrocidades y los excesos de la Revolución Francesa son sólo un ejemplo de los horrores que se han perpetrado en nombre de la libertad (específicamente en nombre de la libertad positiva). La profunda desconfianza que a Berlin le inspiraba la libertad positiva surgía de los excesos del siglo XX, especialmente de los de Stalin. El problema provenía de la creencia (el vicio del reformador social) de que existe un único horizonte deseable para la sociedad, un único remedio para sus males, Contra esta concepción, Berlin se erigió en un firme defensor del pluralismo de los valores humanos. Así, argumentaba que existe una pluralidad de bienes distintos e incompatibles, que nos obligan a tomar decisiones radicales. Su compromiso liberal con la libertad negativa estaba reforzado por su idea de que este tipo de libertad fomentaba el entorno más propicio para que los individuos controlaran y dieran forma a sus vidas a partir de tales elecciones.

En parte, este concepto del yo dividido, que a Berlin le parecía implícito en el concepto de libertad positiva, le inspiraba algunas reservas. Volvamos a George: la parte de él que entiende adecuadamente lo que más le conviene es supuestamente la más elevada, la del yo más racional. Si es incapaz de alentar el predominio de esta parte, tal vez necesite ayuda externa, ayuda de personas más prudentes que él y más capaces de saber cómo debería actuar. Si esto es así, estamos a un paso de considerar legítimo impedir físicamente a George que se acerque a la botella de coñac. Y, teme Berlin, lo que vale para George vale también para el Estado: amparándose en la pancarta de la libertad (positiva) un gobierno puede volverse tiránico y establecer un fin particular para la sociedad; o también priorizar un determinado modo de vida para sus ciudadanos; o incluso decidir lo que deberían desear sin atender a sus deseos reales (véase el cuadro).

«Manipular a los hombres, impelerlos a metas que uno —el reformador social— ve, pero que ellos no pueden ver, es negarles su esencia humana, tratarlos como objetos desprovistos de voluntad propia y, en consecuencia, degradarlos.»

Isaiah Berlin, 1959

La idea en síntesis: libertades en conflicto

45 El principio de la diferencia

La dinámica de las sociedades humanas es inmensamente compleja, pero es razonable suponer que, en general, las sociedades justas son más estables y duraderas que las injustas. Los miembros de una sociedad deben considerar que es, con todo, justa, para poder acatar las leyes que la mantienen cohesionada, y para mantener sus instituciones. Pero ¿cómo deberían distribuirse las obligaciones y los derechos entre los miembros de una sociedad para que fuera justa?

Debemos suponer que la única distribución verdaderamente justa de los bienes sociales es la equitativa. Sin embargo, la equidad puede significar distintas cosas. ¿Significa igualdad de recursos, de tal modo que todo el mundo tiene una participación semejante en la riqueza y los beneficios que la sociedad puede ofrecer, y que todo el mundo debe cargar igualmente con un peso semejante de obligaciones? Pero la espalda de algunos es más grande o más fuerte que la de otros, y la sociedad en su conjunto debe aprovechar los grandes esfuerzos que algunos de sus miembros son capaces de hacer. Si existen personas dispuestas a hacer esfuerzos adicionales, ¿no es razonable que tengan una mayor participación en los beneficios? De lo contrario, quienes están dotados de mayores talentos naturales no podrían explotar plenamente sus dones, y la sociedad en su conjunto se vería perjudicada. De modo que, tal vez, lo importante sea la igualdad de *oportunidades*, por la que cualquier miembro de una sociedad tiene las mismas oportunidades de prosperar, incluso si algunas personas las aprovechan mejor que otras y acumulan más beneficios al hacerlo.

Cronología

1651
El Leviatán

Rawlsianos contra utilitaristas

Buena parte de la fuerza de la idea rawlsiana de la justicia procede de su oposición a una concepción clásica utilitarista sobre los mismos asuntos (véase la página 73). Desde la perspectiva utilitarista, cualquier desigualdad está justificada siempre que revierta en un claro incremento de la utilidad (como en el caso de la felicidad). Así, por ejemplo, los intereses de la mayoría podrían sacrificarse a cambio de un gran beneficio para una minoría; o un perjuicio considerable para una minoría podría justificarse siempre que produjese un beneficio suficiente para una gran mayoría. El principio de la diferencia de Rawls, que

desaconseja sacrificar los intereses de los más desfavorecidos, anula ambas posibilidades.

Otra diferencia importante es que los utilitaristas son imparciales en la consideración de los intereses de cada cual; en efecto, cada cual debe compartir sus propios intereses con los de los demás, y procurar cualquier resultado que pueda proporcionar el mayor beneficio neto en utilidad. Por el contrario, los rawlsianos, que se remiten a la posición original, actúan egoístamente; son el interés propio combinado con la incertidumbre acerca de su futura posición en la sociedad los que explican la aprobación prudencial del principio de la diferencia.

En su *Teoría de la justicia*, publicada en 1971, el filósofo norteamericano John Rawls hizo una contribución decisiva al debate en torno a la justicia social y a la equidad. En el seno de su teoría yace el llamado « principio de la diferencia» , según el cual las desigualdades sociales sólo son justificables si resultan mejores para sus miembros más desfavorecidos de lo que hubiera ocurrido en otro caso. El principio de Rawls suscitó un buen número de críticas, favorables y desfavorables, y se ha invocado, aunque no siempre como al propio Rawls le hubiera gustado, para sostener algunas de las posiciones ideológicas del espectro político.

1971	1974
El principio de la diferencia	La máquina de la experiencia La tierra como bote salvavidas

La teoría del caballo y el gorrión

El principio de la diferencia de Rawls estipula la igualdad excepto cuando la desigualdad beneficia a todos, e impide así que los intereses de un grupo sean subordinados a los de otros. Sin embargo, el principio no dice nada de los beneficios relativos de los distintos beneficiarios, de modo que una pequeña mejora para los más desfavorecidos justificaría una considerable pérdida para quienes ya disfrutan de la mejor parte de los bienes de la sociedad. Esto ha hecho que se invoque este principio desde posiciones muy distantes de la de Rawls, esencialmente igualitarista. Así, la llamada « economía del goteo» de las administraciones de Reagan y Thatcher en los años ochenta se apropió del espíritu rawlsiano al defender que los recortes de los impuestos de los más ricos favorecían el incremento de las inversiones y el crecimiento económico, y en

consecuencia mejorarían las condiciones económicas de los menos favorecidos. Esta reivindicación fue descrita desdeñosamente por J. K. Galbraith como « la economía del caballo y el gorrión» : « tú engorda al caballo, que ya le dará de comer al gorrión alguien que pase por ahí» .

Detrás del velo de la ignorancia Cualquier concepción de la justicia social supone, al menos implícitamente, la noción de la imparcialidad. Un solo indicio de que los principios y las estructuras en las que se basa un sistema social favorecen a un sector particular (una clase social o una casta, por ejemplo, o un partido político), automáticamente convierte en injusto ese sistema. Para entender esta idea de la imparcialidad y fundamentar sus principios de justicia en la equidad, Rawls plantea un experimento mental que tiene su origen en las teorías del contrato social de Rousseau y Hobbes (véase la página 188). Se nos invita a imaginamos a nosotros mismos en lo que él denomina la « posición original» , en la que todavía no existe ningún interés personal o lealtad: « Nadie sabe qué posición ocupa en la sociedad, ni su clase social, ni su estatus, ni siquiera sabe qué dones o habilidades le ha otorgado la naturaleza, qué inteligencia, qué fuerza física, ni otras cosas similares» . Aunque podríamos pretender anteponer nuestros intereses, no sabríamos dónde se encuentran, de modo que no sería posible defender intereses particulares. Al ignorar qué papel tendremos en la sociedad, estamos obligados a ir sobre seguro, y a aseguramos de que no se desfavorece a ningún grupo para favorecer a otro.

«Una sociedad que ponga la igualdad —en el sentido de igualdad de recursos— por encima de la libertad no conseguirá ni la igualdad ni la libertad.»

Milton Friedman, 1980

«El principio de la diferencia constituye una firme concepción igualitarista en el sentido de que a menos que exista una distribución que mejore la situación de ambas personas … es preferible una distribución equitativa.»

John Rawls, 1971

Así, la imparcialidad es —en una paradoja que es tan sólo aparente— la decisión racional e inevitable de sujetos que velan por sus propios intereses en una posición original. Rawls reivindica que las estructuras y los acuerdos sociales y económicos sólo podrían considerarse inequívocamente justos si pudieran suscribirse en este contexto imaginario del « velo de ignorancia» . Además, sólo lo que se acordara en tales circunstancias puede ser, a su vez, lo que estaría dispuesto a aceptar cualquier individuo que actuase racional y prudencialmente.

Y lo mejor y más prudente que puede hacer quien decide racionalmente para salvaguardar sus futuros (inciertos) intereses es atenerse al principio de la diferencia.

El corolario del principio de la diferencia —la idea de que las desigualdades son aceptables sólo si benefician a los más desfavorecidos— es que bajo cualesquiera otras circunstancias las desigualdades son inaceptables. Dicho de otro modo: las condiciones de igualdad deben existir excepto cuando el principio de la diferencia indica que una determinada desigualdad es permisible. Así, por ejemplo, las medidas económicas que mejoran considerablemente las condiciones de los más favorecidos, pero mantienen las de los más desfavorecidos, no se consideran justas. Es posible que haya gente que tiene, por el azar del nacimiento, mayores talentos naturales que otros, pero éstos sólo pueden reportarles mayores ventajas sociales o económicas si con ello contribuyen a una mejora de la situación de los más desfavorecidos. En resumen, la desigualdad sólo es justa si beneficia a todos; de lo contrario debe prevalecer la igualdad.

La idea en síntesis: la justicia como equidad

«Mientras los hombres viven sin un poder común que los atemorice, se hallan en la condición que se denomina estado de guerra; una guerra tal que enfrenta a todos contra todos … En una condición semejante la industria no tiene ninguna oportunidad, ya que su fruto es incierto; por consiguiente no hay agricultura, ni navegación, ni ninguno de los artículos que pueden importarse por mar, ni construcciones confortables, ni instrumentos para mover y desplazar las cosas que requieren mucha fuerza, ni conocimiento de la faz de la tierra, ni cómputo del tiempo, ni artes, ni letras, ni sociedad; y lo peor de todo es que existe un constante temor y peligro de muerte violenta, y la vida del hombre es así solitaria, pobre, tosca, embrutecida y breve.»

«Señalo en primer lugar como inclinación de la humanidad entera, un perpetuo e incesante afán de Poder, que sólo cesa con la muerte.»

Thomas Hobbes, 1951

El pasaje más famoso de una obra maestra de la filosofía política, esta distopía de la humanidad, la presenta el filósofo inglés Thomas Hobbes en su libro *Leviatán*, publicado en 1651. Desalentado por las secuelas inmediatas de la guerra civil inglesa, Hobbes ofrece un panorama de la humanidad que es consecuentemente pesimista y lóbrego: una visión de los humanos, en un imaginario « estado de naturaleza» , aislados, preocupados únicamente por sus propios intereses, cuya única ocupación es su propia seguridad y su propio placer; en constante competición y conflicto recíproco, preocupados sólo por ser los primeros en tomar represalias; entre los cuales no existe la confianza y, por consiguiente, ninguna posibilidad de cooperación.

Cronología

1651	1789
El Leviatán	Teorías del castigo

Contratos sociales

La idea de concebir un contrato social para entender el funcionamiento de un Estado ha seducido a muchos filósofos desde Hobbes. Participar de un contrato confiere a las partes determinados derechos e impone determinadas obligaciones; cabe suponer que una forma paralela de justificación subyace al sistema de derechos y obligaciones que existen entre los ciudadanos de un Estado y las autoridades que lo controlan. Pero exactamente ¿qué tipo de contrato supone o implica esto? El contrato entre el ciudadano y el Estado no tiene un significado literal, y el

« estado de naturaleza» que se imagina en ausencia de sociedad civil es asimismo hipotético, propuesto como un recurso para distinguir entre los aspectos naturales y los convencionales de la condición humana. Pero entonces cabe preguntarse —como hizo el filósofo escocés David Hume—, qué valor se debe atribuir a tales nociones hipotéticas en el momento de determinar los poderes y las prerrogativas concretas del ciudadano y del Estado.

El sucesor más influyente de Hobbes fue el filósofo francés Jean-Jacques Rousseau, cuya obra *El contrato social* se publicó en 1762. Desde entonces ha habido numerosos teóricos del contrato social (o « contractualistas»), el más conspicuo de los cuales es el filósofo norteamericano John Rawls (véase la página 185).

La pregunta para Hobbes es cómo individuos hundidos en semejante estado de discordia espantosa e implacable lograron salir jamás de él. Lo cual es tanto como preguntar: ¿acaso puede desarrollarse alguna forma de organización social y política a partir de semejantes orígenes atroces? He aquí su respuesta: gracias a « un poder común que los atemorice» ; el poder absoluto del Estado, denominado simbólicamente Leviatán.

El buen salvaje

Rousseau, el sucesor francés de Hobbes, no comparte el lóbrego retrato hobbesiano de los humanos en « estado de naturaleza» (es decir, sin las limitaciones que imponen las convenciones sociales y legales). Mientras que Hobbes considera el poder como un medio necesario para domesticar la naturaleza bestial de los individuos, Rousseau piensa que los vicios y otros males de los hombres son el producto de la sociedad: al « buen salvaje» , inocente por naturaleza, feliz en su « razón dormida» , y que vive en armonía con sus semejantes, lo corrompe la educación y otras influencias sociales. La concepción de la inocencia perdida y de los sentimientos no intelectualizados inspiró al movimiento del Romanticismo que se extendió por Europa hacia finales del siglo XVIII. Sin embargo, el propio Rousseau nunca tuvo la menor esperanza de que fuera posible un retorno a cualquier condición idílica anterior: una vez abandonada definitivamente la inocencia, era inevitable el advenimiento del tipo de

constricciones planteadas por Hobbes.

«Sin la espada, los pactos son sólo palabras» Para Hobbes, el instinto de cada individuo es velar por su propio interés, y en atención al propio interés de cada cual conviene cooperar: sólo de este modo puede escaparse del estado de guerra y de la vida « solitaria, pobre, tosca, embrutecida y breve» . Si esto es así, ¿por qué no resulta fácil el acuerdo y la cooperación entre los individuos en el estado de naturaleza? No resulta fácil porque contraer un contrato siempre tiene un precio y no hacerlo siempre supone alguna ventaja (por lo menos a corto plazo). Pero si el interés propio y la autopreservación constituyen la única pauta moral, ¿cómo es posible estar seguros de que los demás no buscarán preventivamente sacar alguna ventaja traicionando el contrato? No cabe ninguna duda de que buscarán sacar ventaja, entonces ¿acaso no es mejor ser el primero en quebrantar el contrato? Naturalmente todos los demás razonan del mismo modo, por lo que no es posible la confianza ni, por lo tanto, el acuerdo. En el estado de naturaleza de Hobbes, los intereses a largo plazo siempre se supeditan a los beneficios a corto plazo, inhibiendo la salida del ciclo de desconfianza y violencia.

«**El hombre nace libre; pero se encuentra encadenado por todas partes. Alguno se cree el señor de los demás, aunque sea más esclavo que ellos. ¿Cómo ha podido esto llegar a producirse? No lo sé. ¿Qué puede legitimarlo? A eso sí creo poder dar respuesta.**»

Jean-Jacques Rousseau, 1782

De bestias y monstruos

Leviatán, que a menudo se asocia con Behemoth, es un aterrador monstruo marino mítico que aparece en varios pasajes relacionados con la creación en el Antiguo Testamento así como en otras fuentes literarias.

Hobbes usa este nombre para sugerir el colosal poder del Estado (« ese gran LEVIATÁN, o mejor dicho —hablando con mayor reverencia— aquel Dios Mortal, sometido al Dios Inmortal, al que debemos nuestra paz y nuestra defensa»).

En el uso moderno, la palabra se suele aplicar al Estado para una apropiación del sugerir poder y de la autoridad más allá de lo debido.

«Sin la espada, los pactos son sólo palabras», concluye Hobbes. Lo que se requiere es alguna forma de poder externo o de sanción que *obligue* a los individuos a acatar los términos del contrato que beneficia a todos (siempre que todos lo acaten). Los individuos deben restringir de buen grado sus libertades en beneficio de la cooperación y de la paz, a condición de que cada cual haga lo propio; deben «otorgar todo su poder y su fuerza a un Hombre, o a una Asamblea de hombres, capaz de reducir todas sus Voluntades, gracias a la pluralidad de voces, a una sola Voluntad». Así es como los ciudadanos acuerdan ceder su soberanía al Estado, que dispone de poder absoluto para «conformar las voluntades de todos ellos con vistas a la Paz en su propio país y a la mutua ayuda contra sus enemigos extranjeros».

La idea en síntesis: el contrato social

47 El dilema del prisionero

«"Éste es el trato: exculparte y testificar contra tu compañero (le caerán 10 años de cárcel y tú simplemente quedarás libre)." Gordon sabía que en cualquier caso la policía podía meterlos en la cárcel un año, sólo por ir armados con cuchillos; pero les faltaban pruebas para imputarles el robo. La trampa era que también sabía que le estaban proponiendo lo mismo a Tony en la celda contigua: si los dos confesaban y se incriminaban mutuamente a cada uno le caerían 5 años. Ojalá pudiera saber qué iba a hacer Tony…

» … Como Gordon no es tonto, sopesa con cuidado las opciones. "Supongamos que Tony no dice nada; en ese caso mi mejor jugada es delatarlo: le caerán 10 años y yo seré libre. Pero supongamos que me acusa a mí: entonces mejor que confesar, es acusarle a él, y que me caigan cinco años (si no, si no digo nada, me caerán los 10 años). De modo que en cualquier caso, tanto si Tony confiesa como si no, la mejor jugada es confesar." Pero el problema de Gordon es que Tony tampoco es tonto y llega exactamente a la misma conclusión. De modo que se incriminan mutuamente y les caen cinco años. Y sin embargo, si ninguno hubiera dicho nada, sólo les hubiera caído un año a cada uno…»

Los dos han tomado una decisión racional, basada en un cálculo de sus propios intereses, pero incluso así el resultado no es el mejor posible para cada uno de ellos. ¿Qué se torció?

Suma cero

La teoría de los juegos ha dado lugar a un campo tan fértil que parte de su terminología ya es de uso común. Un «juego de suma cero», por ejemplo (se suele usar de manera informal, en particular en el mundo de los negocios), es técnicamente un juego como el póquer o el ajedrez, en el que las ganancias de un lado quedan exactamente compensadas por las pérdidas del otro, de modo que la suma de ambas es cero. En cambio, el juego del prisionero no es de «suma cero», pues cabe la posibilidad de que ambos jugadores ganen… o pierdan.

Cronología

c. 350 a. C.	1789
Formas del razonamiento	Teorías del castigo

La teoría de juegos Este resumen del relato conocido como el «dilema del

prisionero» posiblemente sea el más célebre de una serie de situaciones estudiadas en el campo de la teoría de juegos. El objeto de la teoría es analizar situaciones de este tipo, donde existe un claro conflicto de intereses, y determinar en qué consistiría una estrategia racional. Tal estrategia, en este contexto, es la que permite maximizar las propias ventajas e implica, o bien aliarse con un adversario (la «cooperación» en los términos de la teoría de juegos), o bien traicionarlo (la «deserción»). Naturalmente, se supone que tal análisis arroja luz en el comportamiento humano real, explicando por qué la gente actúa como lo hace, o bien prescribiendo como debería hacerlo.

En el análisis de la teoría de juegos, las estrategias posibles que pueden adoptar Gordon y Tony pueden presentarse en una «matriz de pagos»:

	Tony no dice nada	**Tony confiesa**
Gordon no dice nada	Ambos cumplen 1 año (gana-gana)	Gordon cumple 10 años Tony queda en libertad (pierde mucho-gana mucho)
Gordon confiesa	Gordon queda en libertad Tony cumple 10 años (gana mucho-pierde mucho)	Los dos cumplen 5 años (pierde-pierde)

1950 El dilema del prisionero
1976 ¿Es malo tener mala suerte?

El dilema se produce porque la preocupación exclusiva de cada prisionero es minimizar los años de prisión. Para lograr el mejor resultado para ambos colectivamente (un año para cada uno), deberían colaborar y ponerse de acuerdo en renunciar al mejor resultado individual (la libertad). En el planteamiento clásico del dilema no existe la posibilidad de tal colaboración, y en cualquier caso ninguno de los dos prisioneros tendría razón alguna para confiar en

que el otro no incumpla el acuerdo. De modo que adoptan una estrategia que excluye el mejor resultado colectivo, para evitar el peor resultado individual, y van a parar a un resultado intermedio que no es el más óptimo.

Implicaciones en el mundo real Las implicaciones generales del dilema del prisionero son que la persecución egoísta de los propios intereses, aun siendo racional en ocasiones, no puede conducirnos al mejor resultado para nosotros y para los demás; y además que la colaboración, al menos en determinadas circunstancias, es globalmente la mejor política. Pero ¿qué ocurre si trasladamos el dilema del prisionero al mundo real?

El dilema del prisionero ha tenido un influjo especial en las ciencias sociales, sobre todo en la economía y la política. Puede iluminar, por ejemplo, la toma de decisiones y la psicología en la que descansan las escaladas armamentísticas entre naciones enfrentadas. En tales situaciones, en principio resulta claramente beneficioso para las dos partes implicadas llegar a un acuerdo para limitar el nivel de gasto armamentístico, pero en la práctica en muy raras ocasiones ocurre así.

De acuerdo con el análisis de la teoría de juegos, no se llega a alcanzar acuerdos por el temor a una pérdida grande (una derrota militar) mucho peor que una ganancia relativamente pequeña (un menor gasto militar); el resultado real —ni el peor ni el mejor posibles— es la carrera armamentística.

Se ha visto un claro paralelismo entre el dilema del prisionero y el sistema que consiste en declararse culpable de un delito menor para evitar la acusación de un delito más grave y que se da en algunos sistemas judiciales (en Estados Unidos) pero que está prohibido en otros. La lógica del dilema del prisionero sugiere que la estrategia racional de «minimizar las pérdidas mayores» —es decir, aceptar una sentencia o una pena más leve por temor a recibir una más grave— puede inducir a las partes inocentes a confesar o inculparse mutuamente. En el peor de los casos, ello podría conducir a que la parte culpable confesara su falta mientras la inocente siguiera alegando su inocencia, con la extraña consecuencia de que el inocente recibiera la pena más severa.

Una mente maravillosa

El teórico actualmente más famoso de la teoría de juegos es John Forbes Nash de Princeton. Su genio matemático y su triunfo contra la enfermedad mental le valieron un Premio Nobel de Economía en 1994, lo cual inspiró la película *Una mente maravillosa* (2001).

Como teórico de los juegos, la contribución más célebre de Nash es el

haber definido el epónimo « equilibrio de Nash» : una situación estable en un juego en la que ningún jugador tiene ningún incentivo para cambiar su estrategia, a menos que otro jugador cambie la suya. En el dilema del prisionero, la doble delación (los dos jugadores confiesan) representa el equilibrio de Nash, que, como se ha visto, no se corresponde necesariamente con el resultado óptimo para los jugadores implicados.

Gallinas Otro juego muy estudiado por los teóricos es el del « gallina» , cuya versión más memorable se encuentra en la película *Rebelde sin causa* que interpretó James Dean en 1955. Aquí, dos jugadores conducen sus respectivos coches uno contra el otro y el perdedor (el gallina) es el que se desvía para evitar chocar. En este escenario, el precio de la cooperación (desviarse y quedar como un gallina) es tan pequeño en comparación con el precio de la delación (seguir recto y chocar) que lo racional en este caso parece ser la colaboración. El peligro surge cuando el jugador A supone que el jugador B es igualmente racional y que, en consecuencia, se desviará, lo cual le permite a él (el jugador A) seguir impunemente en línea recta y ganar.

El peligro inherente en este juego es obvio: la doble delación (los dos siguen en línea recta) implica el choque seguro. Los paralelismos con varios tipos de políticas arriesgadas del mundo real (la más potencialmente peligrosa de las cuales es la nuclear) son también claros.

La idea en síntesis: jugar en serio

El signo de una sociedad civilizada, dirían muchos, es su capacidad de defender los derechos de sus ciudadanos: de defenderlos contra un trato arbitrario o perjudicial por parte del Estado o de otros individuos, para permitir su plena expresión política y garantizar la libertad de expresión y de movimiento. Entonces ¿por qué tal sociedad debería infligir daños a sus ciudadanos deliberadamente, excluirlos de sus procesos políticos, restringir su libertad de movimiento o de expresión? Ésa es exactamente la prerrogativa que el Estado se toma cuando decide castigar a sus ciudadanos por vulnerar las reglas que les ha impuesto.

Este conflicto aparente entre diferentes funciones del Estado configura el debate filosófico sobre la justificación del castigo. Como ocurría en la discusión de otros asuntos éticos, el debate sobre la justificación del castigo ha tendido a dividirse entre planteamientos consecuencialistas y deontológicos (véase la página 69): las teorías consecuencialistas hacen hincapié en las consecuencias beneficiosas del castigo a los infractores; mientras que las teorías deontológicas insisten en que el castigo es intrínsecamente bueno como un fin en sí mismo, con independencia de cualesquiera otros beneficios que pueda comportar.

«Es exactamente lo que merecen» La idea clave de las teorías que sostienen que el castigo es bueno en sí mismo es la reparación. Una convicción básica que subyace a la mayor parte de nuestras ideas morales es que debería darse a la gente lo que se merece: del mismo modo que se beneficiarían si se portasen bien, deben sufrir por portarse mal. La idea de la reparación (la gente debe pagar un precio —a saber, la pérdida de la libertad— por su mal comportamiento) condice perfectamente con esta convicción. En ocasiones se introduce una idea adicional: la noción de que la infracción crea un desequilibrio, y de que el equilibrio moral se restaura haciendo que el infractor «pague su deuda» a la sociedad; el delincuente ha asumido la obligación con la sociedad de no quebrantar las reglas, de modo que al quebrantarlas contrae una pena (una deuda o un cargo) que debe pagar. La metáfora financiera puede extenderse perfectamente y demandarse una transacción justa: que la severidad de la pena se adecúe a la severidad del crimen.

Cronología

c. 400 a. C.	1651	1785
La defensa del libre albedrío	El Leviatán	Medios y fines

Niveles de justificación

El «problema del castigo» a menudo adopta como justificaciones consideraciones de tipo utilitarista como la disuasión, la protección de la sociedad y/o factores intrínsecos como la reparación. Pero también puede implicar cuestiones más generales o más específicas. En un nivel específico, podemos preguntarnos si el castigo de un individuo *particular* está justificado. Esta pregunta no pone en duda la propiedad general del castigo y no es de un interés filosófico específico o exclusivo. Estas preguntas tienen implicaciones en términos de responsabilidad. ¿Era el acusado responsable de sus acciones en el sentido requerido por la ley? ¿O estaba actuando bajo alguna coacción o en defensa propia? Aquí, la cuestión de la responsabilidad nos lleva a algunos terrenos filosóficos muy complicados. En el nivel más general, el problema del libre albedrío plantea una pregunta: si todas nuestras acciones están predeterminadas, ¿ejercemos la libertad de decisión en cada uno de nuestros actos y, si no es así, podemos ser responsables de *algo* de lo que hacemos?

La idea de que «el castigo debería ajustarse al crimen» encuentra apoyo en la *lex talionis* (la «ley del tabón») del Antiguo Testamento: «ojo por ojo diente por diente». Ello implica que el crimen y el castigo deben ser equivalentes no sólo en severidad sino en naturaleza. Los defensores de la pena de muerte, por ejemplo, suelen alegar que la única reparación cabal del asesinato es la pérdida de la vida (véase la página 198); aunque quienes piensan así no puedan proponer que los chantajistas sean chantajeados o que los violadores sean violados. Este fundamento bíblico de las teorías de la reparación plantea el principal problema con el que se enfrentan; pues la *lex talionis* es obra de un Dios «vengativo», y el reto de la teoría de la reparación es mantener una distancia respetable entre la venganza y la reparación. La idea de que algunos crímenes «claman» castigo se disfraza a veces diciendo que el castigo expresa la indignación de la sociedad hacia un determinado acto porque, cuando la reparación queda puesta en evidencia como poco más que afán de venganza, parece muy poco adecuada como justificación del castigo.

La pena de muerte

Los debates sobre la pena de muerte normalmente se estructuran de un modo parecido a los debates sobre otros tipos de castigo. Los defensores de la pena capital suelen argumentar que es justo castigar los crímenes más serios con la pena más severa, con independencia de las consecuencias beneficiosas que pueda comportar; pero los supuestos beneficios —sobre todo la disuasión y la protección del pueblo— a menudo también son esgrimidos. Los que se oponen replican señalando que el valor disuasorio es dudoso en el mejor de los casos, que el encarcelamiento brinda la misma protección al pueblo, y que la instauración de la pena capital envilece a la sociedad. El argumento de más peso contra la pena capital —la certeza de que a veces se ha ejecutado a inocentes y de que se los seguirá ejecutando— es difícil de refutar. Tal vez, el mejor argumento a favor de la pena de muerte sea que la muerte es preferible a, o menos cruel que, una vida entre rejas, pero esto sólo nos serviría a lo sumo para concluir que debería dejarse decidir al delincuente si prefiere morir o vivir.

«**Todo castigo es un delito: todo castigo es en sí mismo malvado.**»

Jeremy Bentham, 1789

El mal necesario En las antípodas de las posiciones de la reparación, las justificaciones utilitaristas, u otras de tipo consecuencialista, del castigo se caracterizan no sólo por negar que sea bueno sino por considerarlo como positivamente malo. El pionero del utilitarismo clásico, Jeremy Bentham, pensaba que el castigo era un mal necesario: era malo porque aumentaba la cantidad de desdicha humana; y sólo se justificaba en la medida en que los beneficios que supusiera superasen la desdicha que causaba. No era ésta una posición puramente teórica, tal y como dejara claro la eminentemente práctica reformista de las cárceles del siglo XIX, Elizabeth Fry: « El castigo no es una venganza sino un modo de reducir el crimen y de reformar al criminal» .

Habitualmente, se entiende que el castigo como instrumento para reducir el crimen satisface su papel de dos formas: la inhabilitación y la disuasión. Sin duda, un asesino ejecutado no podrá volver a matar, pero tampoco podrá hacerlo uno que esté encarcelado. El grado de inhabilitación —especialmente en la inhabilitación permanente por medio de la pena capital— puede suscitar discusiones, pero la necesidad de medidas de este tipo en beneficio del interés público es difícil de discutir. Los argumentos en tomo a la disuasión son menos claros. Parece perverso decir que alguien debería ser castigado, no por el crimen

cometido sino para disuadir a otros de cometer ofensas similares; y también su utilidad práctica plantea dudas, pues los estudios sugieren que lo disuasorio es más bien el miedo a ser capturado y no cualquier castigo derivado.

Otro aspecto fundamental en el pensamiento utilitarista sobre el castigo es la reforma o la rehabilitación del criminal. Para el individuo con mentalidad liberal resulta muy seductora la idea de que el castigo constituya una forma de terapia mediante la que se reeduca y se reforma a los delincuentes de modo que puedan convertirse en plenos y útiles miembros de la sociedad. Sin embargo, existen dudas serias sobre la capacidad de los sistemas penales —los habituales, al menos — para conseguir este propósito.

En la práctica, es fácil ofrecer contraejemplos que muestran la inadecuación de cualquier justificación utilitaria particular del castigo: basta con citar casos en los que el delincuente no representa un peligro para la ciudadanía o no necesita reforma, o en los que el castigo carece de cualquier valor disuasorio. Sin embargo, el planteamiento habitual consiste en ofrecer una colección de posibles beneficios que el castigo brinda, sin sugerir que todos ellos se aplican en todos los casos. Pero incluso así, cabe la posibilidad de que sigamos sintiendo que a la explicación puramente utilitaria le falta algo, y que debe preservarse algún espacio para la reparación. De acuerdo con este sentimiento, muchas teorías recientes son de naturaleza esencialmente híbrida, al intentar combinar elementos de la teoría de la reparación con otros de la utilitarista para ofrecer una explicación global del castigo. Entonces, la principal labor consistiría en establecer prioridades en los distintos objetivos especificados, y poner de relieve los puntos en los que entran en conflicto con las actuales políticas y prácticas.

La idea en síntesis: ¿se adecúa el castigo al crimen?

49 La Tierra como bote salvavidas

«*A la deriva en un Mar Moral* … Digamos que aquí estamos sentadas 50 personas, en nuestro bote salvavidas. Siendo generosos, pongamos que caben 10 más, en total 60. Supongamos que los que vamos en el bote salvavidas vemos a 100 más que nadan en el agua, implorando que los dejemos subir a nuestro bote o pidiendo limosna…

» … Tenemos muchas opciones: podemos estar tentados de intentar vivir de acuerdo con el ideal cristiano de “cuidar de nuestros hermanos”, o con el ideal marxista de vivir “cada cual de acuerdo con las propias necesidades”. Pero puesto que las necesidades de todos en el agua son las mismas, y puesto que todos pueden considerarse nuestros hermanos, podríamos ofrecer nuestro bote a todos, con lo cual sumaríamos 150 en un bote diseñado para 60. El bote se hunde, y se ahoga todo el mundo. Una justicia completa y una catástrofe completa … Puesto que el bote tiene capacidad para 10 pasajeros más, podríamos admitir solamente a 10. Pero ¿a cuáles dejaríamos subir? Supongamos que decidimos no admitir a ninguno. Así podríamos sobrevivir aunque deberíamos estar vigilando constantemente a los que quieren abordarnos.»

En un artículo publicado en 1974 el ecologista norteamericano Garrett Hardin introdujo la metáfora de la Tierra como bote salvavidas para argumentar contra las ayudas de los países ricos occidentales a los países en vías de desarrollo de todo el mundo. Erigido en el incansable azote de la sensiblería liberal, Hardin sostenía que las intervenciones bienintencionadas pero equivocadas de Occidente estaban perjudicando a largo plazo a todos los implicados. Los países que recibían ayuda extranjera desarrollaban una cultura dependiente y así no conseguían « aprender a largo plazo» los peligros de un plan posterior inadecuado y del crecimiento de la población ilimitado. Al mismo tiempo, la inmigración sin restricciones significa que las casi estancadas poblaciones occidentales se verán ahogadas rápidamente por las ininterrumpidas oleadas de refugiados económicos. Hardin atribuye la culpa de estos males a la aterrorizada conciencia de los liberales, y censura especialmente el que alentaran la « tragedia de los comunes» , un proceso en el que los limitados recursos, considerados de modo idealista como propiedad de todos los humanos, quedarían incluidos en un tipo de propiedad común que inevitablemente conducía a la sobreexplotación y a la ruina (véase el cuadro).

Cronología

c. 30 d. C.	1959
La regla de oro	Libertad positiva y negativa

La tragedia de los recursos comunes

El recurso de Hardin a la dura ética del bote salvavidas era una respuesta directa a los problemas que había observado en la amable metáfora de la «nave espacial Tierra» tan cara a los ecologistas soñadores; de acuerdo con ella, todos nosotros viajamos a bordo de una nave espacial, de modo que tenemos el deber de asegurarnos que no derrochamos los preciados y limitados recursos de la nave. El problema se plantea cuando este panorama se funde con la amable tripulación en la que todos están bien avenidos, y que alimenta la concepción de que los recursos mundiales deberían gestionarse en común, y todo el mundo debería recibir una parte justa e igual de los mismos. Un granjero que tiene una parcela de tierra se ocupará de su propiedad y se asegurará de no arruinarla sobreexplotándola, pero si se convierte en terreno común abierto a todos, no estará tan preocupado en protegerla. Las tentaciones de la ganancia a corto plazo significan que las restricciones voluntarias desaparecerían pronto, y la degradación y el declive las sucederían rápidamente. Estos procesos —inevitables, según Hardin, en «un mundo abarrotado de seres humanos que están bastante lejos de ser imagen liberal de una gran y feliz perfectos»— constituyen lo que él llama la «tragedia de los comunes». En esta exacta medida, cuando los recursos de la Tierra, como el aire, el agua o los peces de los océanos, son tratados como recursos comunes, no existe una administración adecuada de los mismos y la ruina está asegurada.

1971 El principio de la diferencia

1974 La Tierra como bote salvavidas

«La ruina es el destino hacia el que todos los hombres se apresuran, pues cada uno persigue su propio interés en una sociedad convencida de la libertad para los recursos comunes. La libertad en los recursos comunes conduce a la ruina para todos.»

Garrett Hardin, 1968

La ética del amor incondicional El propio Hardin es poco favorable a promover la ética del amor incondicional en su bote salvavidas. Con la conciencia muy tranquila, su advertencia a la aterrorizada conciencia culposa de los liberales es «largarse y dejar su sitio a los otros», con lo cual se eliminan los remordimientos que amenazan la estabilidad del bote. No hay razón para preocuparse por el modo en que hemos llegado al bote: «no podemos deshacer el pasado»; y sólo cabe aceptar su dureza, la firme postura de que es posible salvar el mundo (o nuestra parte del mundo,

al menos) para las futuras generaciones.

La representación de la relación entre los países ricos y los pobres es realmente horrible: los primeros se encuentran apalancados sin peligro en sus botes y les parten la cabeza en dos o les aplastan los nudillos a los segundos con sus remos cuando intentan trepar a bordo. Pero hay maneras distintas a la de Hardin de interpretar la metáfora. ¿Acaso el peligro de que se hunda el bote salvavidas es real? ¿Cuál es su capacidad real? ¿No se tratará más bien de un problema de que las hinchadas bestias se apretaran un poco y redujeran sus raciones de comida?

Una buena parte del argumento de Hardin descansa en el supuesto de que los elevados índices de natalidad de los países pobres persistirían incluso si tuvieran mejores condiciones; no admite que tales índices pueden ser un *efecto* de una mortalidad infantil más elevada, de una expectativa de vida baja, de una pobre educación, y muchas otras cosas del mismo tipo. El liberal diría que sin el maquillaje lo que nos deja Hardin es la idea de una inmoralidad en bruto y desnuda: el egoísmo, la complacencia, la falta de compasión…

Las fronteras morales Cuando se considera bajo esta luz, la culpa del liberal no parece tan fuera de lugar. Un titular liberal del bote no soñaría en aporrear en la cabeza a un compañero de viaje con el remo, pero ¿qué puede parecerle contemplar cómo otro lo hace (o incluso permitir que suceda algo así) con los desafortunados que están en el agua, alrededor del bote? Además, asumiendo que hubiera efectivamente sitio en el barco, ¿no es su deber ayudarles a salir del agua y compartir las raciones?

El panorama del bote salvavidas confronta así al liberalismo occidental a un reto serio. Uno de los requisitos más elementales de la justicia social es el trato imparcial para todos los individuos; no debería permitirse que las cosas que están más allá de nuestro control (los factores accidentales debidos al nacimiento, por ejemplo, como el color de la piel, el género, etc.) determinaran cómo es tratada y juzgada moralmente esa persona. Y no obstante, existe un factor —el lugar donde nace cada uno— que parece jugar un papel muy importante en nuestras vidas morales, no sólo para los defensores de Hardin, sino también para la mayoría de los que se autoproclaman liberales. ¿Cómo puede tener tanto —siquiera alguno— peso moral algo tan arbitrario como las fronteras nacionales?

Frente a este desafío, el liberal debe, o bien mostrar por qué se suspende el requisito de igualdad, o bien disolver tal requisito cuando consideramos partes del mundo que no son la nuestra (es decir: por qué es correcto para nosotros tener preferencia moral por nosotros mismos); o también puede aceptar que existe alguna incoherencia en el corazón del liberalismo corriente y que la consistencia exige que los principios de la moralidad y de la justicia social se apliquen globalmente.

Teóricos contemporáneos han intentado abordar el asunto de los dos modos. El argumento de la parcialidad es un ingrediente esencial del pensamiento liberal que, si bien es muy útil para hacer frente a realidades globales, sin duda reduce su alcance y su dignidad. Por otro lado, el liberalismo completamente cosmopolita, aunque laudable, exige un cambio radical de las prácticas y las políticas actuales de compromiso internacional que hace temer que se vaya a pique al chocar con esas mismas realidades globales. Sea como fuere, a la filosofía política le queda mucho camino por recorrer en el campo de la justicia internacional y global.

«**La supervivencia futura inmediata exige que actuemos de acuerdo con la ética del bote salvavidas. Le haremos un flaco favor a la posteridad si obramos de otro modo.**»

Garrett Hardin, 1974

La idea en síntesis: ¿queda espacio en el bote?

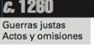
Aunque la guerra siempre ha tenido entusiastas, la mayoría compartiría los sentimientos del poeta Charles Sorley, quien, en 1915, pocos meses antes de morir a los 21 años en la batalla de Loss, escribió: «No existen las guerras justas. Lo que hacemos es combatir a Satán con Satán». Pero casi cualquiera estaría de acuerdo en que, por más que la guerra sea diabólica, algunos diablos son peores que otros. Sí, debería evitarse la guerra, pero no a cualquier precio. Puede ser el menor de los males; el motivo puede ser tan irrenunciable, la causa tan importante, que el recurso a las armas esté moralmente justificado. En tales circunstancias, la guerra puede ser justa.

> «**Nunca ha habido una guerra buena, ni una paz mala.**»
>
> **Benjamin Franklin, 1783**

El debate filosófico sobre la moralidad de la guerra, un asunto a lo sumo tan vigente hoy como siempre, tiene una larga historia. En Occidente, las preguntas surgieron originalmente en las antiguas Grecia y Roma, y de ellas las heredó la Iglesia cristiana. La conversión del Imperio romano al cristianismo en el siglo IV d. C. demandaba un compromiso entre las aspiraciones pacifistas de la temprana iglesia y las necesidades militares de los gobernantes del Imperio. San Agustín impulsó esta reconciliación, que asumió santo Tomás de Aquino, desarrollando la distinción aún hoy canónica entre *ius ad bellum* («justicia en la declaración de guerra», las condiciones bajo las cuales es moralmente correcto tomar las armas) y *ius in bello* («justicia en la guerra», las reglas de conducta una vez empieza la lucha). El debate en la «teoría de la guerra justa» se estructura esencialmente en torno a estas dos ideas.

Las condiciones de la guerra Los principales objetivos de la teoría de la guerra justa son identificar una serie de condiciones bajo las cuales es moralmente defendible recurrir a la fuerza de las armas, y ofrecer pautas sobre los límites para el desarrollo del combate. Los principios de la *ius ad bellum* se han discutido y revisado mucho a lo largo de los siglos. Algunos resultan más controvertidos que otros; en la mayoría, los problemas son de interpretación. Se suele coincidir en que las diversas condiciones son todas necesarias pero no suficientes para justificar una declaración de guerra. Sí se ha conseguido bastante consenso acerca de las siguientes:

Cronología



Causas justas Aún hoy se discute la condición primordial para que una guerra sea moralmente defendible: la causa justa. En siglos anteriores se interpretaba de un modo muy amplio y podía incluir, por ejemplo, la motivación religiosa; en el Occidente secular tales causas se descartarían hoy por considerarse ideológicas, cuando no inadecuadas.

«Bismarck hizo guerras "necesarias" y mató a miles de personas; los idealistas del siglo XX hacen guerras "justas" y matan a millones.»

A. J. P. Taylor, 1906-1990

La mayoría de los teóricos modernos han restringido el alcance de las condiciones a la defensa contra la agresión. Menos controvertido resulta incluir la autodefensa contra la violación de los derechos fundamentales de un país (su soberanía política y su integridad territorial, por ejemplo, Kuwait contra Irak en 1990-1991); y muchos de ellos la ampliarían a la posibilidad de ayudar a un tercero que fuera víctima de una agresión (es el caso de la coalición de fuerzas que liberó Kuwait en 1991). Mucho más controvertidas son las acciones militares preventivas contra potenciales agresores, porque las pruebas definitivas de la intención están necesariamente ausentes. La duda está aquí en si la fuerza preventiva no es en sí misma una agresión, y hay quien considera que sólo una agresión concreta —consumada— puede constituir una causa justa.

Buenas intenciones Están estrechamente vinculadas a la anterior. No basta con tener una causa justa; es necesario que el objetivo único y exclusivo de la acción militar sea favorecer esa causa. Santo Tomás de Aquino hablaba de que debe promoverse el bien y evitarse el mal, pero el asunto crucial es simplemente que la única motivación debería ser restablecer los males provocados mediante la agresión que proporcionó una causa justa. La causa justa no puede ser entonces un pretexto para otros motivos como el interés nacional, la expansión territorial o la exaltación de la grandeza nacional.

1785 Medios y fines

1971 El principio de la diferencia

«Ius in bello»

Otro aspecto de la teoría de las guerras justas es la *ius in bello*: qué es una conducta moralmente aceptable y adecuada una vez empieza el combate. Esto abarca muchas cosas, desde el comportamiento de los soldados en sus relaciones tanto con el enemigo como con los civiles, hasta las principales cuestiones estratégicas, como el uso de armas (nucleares, químicas, minas, bombas de racimo, etc.). En este ámbito, normalmente hay dos consideraciones ineludibles: *La proporción*

requiere que los medios y los fines se adecúen. Por tomar un caso extremo, casi cualquiera acepta que un ataque nuclear no puede justificarse por más útil que pudiera ser para conseguir algún objetivo militar. *La discriminación* requiere que los combatientes y los que no combaten sean rigurosamente diferenciados. Por ejemplo, se considera inadmisible incluir a los civiles, incluso aunque resultara de ayuda para minar la moral de los militares enemigos.

Es evidente que es posible que una guerra justa se desarrolle injustamente, y una injusta de un modo justo. Dicho de otro modo, los requisitos de la *ius ad bellum* y los de la *ius in bello* son distintos, y las dos pueden satisfacerse mutuamente. En particular, muchos aspectos de la *ius in bello* se solapan con el objeto del derecho internacional (tales como el Convenio de La Haya y la Convención de Ginebra), y las infracciones tanto en el bando de los ganadores como en el de los derrotados deben juzgarse en principio como crímenes de guerra.

«**En la guerra no existen vías intermedias.**»

Winston Churchill, 1949

La autoridad competente La decisión de ir a las armas sólo pueden tomarla « las autoridades competentes» de acuerdo con los procesos requeridos. En este contexto « competente» significa cualquier cuerpo o institución del Estado que tenga poder soberano (su competencia para declarar la guerra se suele establecer en el seno de la Constitución de cada país). « Declarar la guerra» es importante, puesto que en ocasiones implica que la decisión de tomar las armas se declarará formalmente a los propios ciudadanos y al Estado(s) contrario. Pero esto parece poco recomendable si el hacerlo concede alguna ventaja estratégica al enemigo, el cual ha perdido efectivamente cualquier derecho a tales consideraciones al iniciar la agresión. El concepto mismo de « autoridad competente» es muy intrincado, y suscita preguntas peliagudas sobre la legitimidad del gobierno y sobre la relación óptima entre los que toman las decisiones y el pueblo.

El último recurso El recurso a la guerra sólo está justificado si —por justa que sea la causa— se han explorado, o cuando menos considerado, cada una de las alternativas pacíficas, no militares. Por ejemplo, si un conflicto puede impedirse por medios diplomáticos, sería categóricamente erróneo dar una respuesta militar. Las sanciones económicas o de otro tipo también deberían considerarse.

Perspectivas de éxito Incluso si se cumplen las anteriores condiciones para una intervención militar, un país sólo debería recurrir a la guerra si tiene alguna posibilidad « razonable» de ganar. Este requisito suena bastante prudente: no tiene

ningún sentido arrasar vidas y recursos en vano. Sin embargo ¿cuán exitoso debe ser el éxito? ¿Es realmente *erróneo* que un poder la emprenda contra un agresor más fuerte con independencia de hasta qué punto las posibilidades de ganar sean desfavorables? El regusto consecuencialista de esta condición resulta ofensivo a muchos. En ocasiones, es perfectamente adecuado resistir a un agresor —e inmoral o cobarde, no hacerlo— por fútil que parezca la empresa.

La proporción Un equilibrio entre la finalidad deseada y las consecuencias que previsiblemente cabe esperar: el bien que se espera (reparar el mal, que constituye la causa justa) debe sopesarse frente a los daños previsibles (las víctimas, el sufrimiento humano, etc.). De modo que cabe exigir que la acción militar ocasione mayores bienes que daños; el beneficio debe ser mayor que el coste. Ésta es también una consideración prudencial y claramente consecuencialista (aunque en este caso casi irresistible SI el bien resultante y el daño pueden definirse y medirse cuidadosamente). Cuando consideramos la proporción entre los medios militares y los fines, empezamos a adentramos en el territorio del *ius in bello* (véase el cuadro de la página anterior).

«**La política es la guerra sin derramamiento de sangre, mientras que la guerra es la política con derramamiento de sangre.**»

Mao Zedong, 1938

No sólo una guerra justa Entre los filósofos contemporáneos, la teoría de la guerra justa tal vez sea el ámbito de mayor debate, aunque no es la única perspectiva. Los dos puntos de vista más extremos son el realismo y el pacifismo. Los realistas son escépticos con respecto al proyecto mismo de aplicar conceptos éticos para pensar en la guerra (o cualquier otro aspecto de la política exterior); la influencia internacional y la seguridad nacional son las principales preocupaciones (los jugadores reales del mundo juegan a béisbol, la moralidad es para los débiles). Los pacifistas, en las antípodas, creen que la moralidad debe prevalecer en los asuntos internacionales; para ellos, la acción militar nunca es la solución correcta: siempre existe una opción mejor.

La idea en síntesis: lucha en un combate limpio

Glosario

Absolutismo En ética, consiste en considerar que determinadas acciones son correctas o incorrectas con independencia de cualquier circunstancia o de sus consecuencias.

Analítico Describe una proposición que aporta mayor información de la que se encuentra contenida en el significado de los términos incluidos. Por ejemplo: «Todos los sementales son machos». En cambio, una proposición que proporciona información significativa («Los sementales son más veloces que las yeguas») se denomina sintética.

Analogía Comparación en la que dos cosas se parecen: un argumento por analogía usa similitudes evidentes entre dos cosas para defender la similitud en otros aspectos no evidentes.

Antirrealismo *véase* Subjetivismo.

A posteriori *véase* A priori.

A priori Describe una proposición cuya verdad puede conocerse sin recurrir a la evidencia de la experiencia. En cambio, una proposición que requiere el recurso a la experiencia es denominada a posteriori.

Consecuencialismo En ética, la noción de que la rectitud de una acción sólo puede establecerse por referencia a la efectividad con que permite alcanzar determinados fines o estados de hecho deseables.

Contingente Describe algo que es de hecho cierto pero podría haber sido de otro modo. En cambio, una verdad necesaria es la que no podría haber sido de otro modo; algo verdadero en cualquier circunstancia o en cualquier mundo posible.

Deducción Una forma de inferencia en la que la conclusión se sigue de (está contenida en) las premisas; si las premisas de un razonamiento deductivo válido son verdad, la conclusión es indudablemente verdad.

Deontología La concepción por la que algunas acciones se consideran intrínsecamente correctas o incorrectas, con independencia de sus consecuencias; se pone especial énfasis en los deberes y las intenciones de los sujetos.

Determinismo Teoría según la cual cada acontecimiento tiene una causa anterior y que, en consecuencia, cada estado del mundo requiere un estado previo que lo determina. El problema del libre albedrío surge cuando el determinismo es llevado hasta el extremo de menoscabar nuestra libertad de acción.

Dualismo En la filosofía de la mente, la noción de que la mente (o el alma) y la materia (o el cuerpo) son distintas. Los dualistas de la sustancia sostienen que la mente y la materia son sustancias esencialmente distintas; los dualistas de las propiedades sostienen que una persona tiene dos tipos de propiedades

esencialmente distintas, las mentales y las físicas. Al dualismo se opone el idealismo o inmaterialismo (las mentes y las ideas son las únicas que existen), y el fisicalismo o el materialismo (los cuerpos y la materia son lo único que existe).

Empírico Describe los conceptos o creencias basados en la experiencia (a saber, los datos de los sentidos o las evidencias de los sentidos); una verdad empírica es la que puede confirmarse como tal apelando únicamente a la experiencia.

Empirismo La concepción de que el conocimiento está basado en, o inextricablemente unido a, la experiencia derivada de los sentidos; se trata de una negación de un conocimiento a priori.

Epistemología Teoría del conocimiento que incluye su fundamentación y justificación, así como el papel de la razón y/o de la experiencia, en su adquisición.

Escepticismo Posición filosófica en la que se desafía nuestra pretensión de conocimiento en uno o en todos los ámbitos.

Estética Es la rama de la filosofía consagrada a las artes, incluida la definición y la naturaleza de las obras de arte, el fundamento del juicio estético, y la justificación de los juicios estéticos y de la crítica de arte.

Falacia Un error de razonamiento. Por lo general se distingue entre las falacias formales, en las que el fallo se debe a la estructura lógica de un argumento, y las falacias informales, que comprenden todas aquellas otras formas en que el razonamiento puede extraviarse.

Fisicalismo *véase* Dualismo.

Idealismo *véase* Dualismo.

Inducción Forma de inferencia en la que se extrae una conclusión empírica (una ley general o un principio) de premisas empíricas (observaciones particulares sobre cómo son las cosas en el mundo); la conclusión sólo se sostiene en las premisas (sin estar incluida en ellas), de modo que las premisas pueden ser ciertas y aun así la conclusión ser falsa.

Inferencia Proceso de razonamiento que va de las premisas a la conclusión; los principales tipos de inferencia son la deducción y la inducción. Discernir entre las inferencias correctas y las incorrectas es el propósito de la lógica.

Inmaterialismo *véase* Dualismo.

Liberalismo (ético) La concepción de que el determinismo es falso y de que las decisiones humanas son genuinamente libres.

Libre albedrío *véase* Determinismo.

Lógica *véase* Inferencia.

Materialismo *véase* Dualismo.

Metafísica Rama de la filosofía que se ocupa de la naturaleza o de la estructura de la realidad, centrándose generalmente en nociones como la de ser, sustancia y causa.

Naturalismo En ética, la concepción de que los conceptos morales pueden explicarse o analizarse meramente en términos de «hechos de naturaleza» que, en principio, puede descubrir la ciencia.

Necesario *véase* Contingente.

Normativo Relativo a las normas (pautas o principios) mediante las que se juzga o se orienta la conducta humana. La distinción normativo/descriptivo es paralela a la distinción entre valores y hechos.

Objetivismo En ética y en estética, la concepción de que valores y propiedades como la bondad y la belleza son inherentes o intrínsecos a los objetos, y existen independientemente de la capacidad humana para aprehenderlos.

Paradoja En lógica, argumento en el que premisas aparentemente inobjetables llevan, mediante un razonamiento aparentemente sólido, a una conclusión inaceptable o contradictoria.

Racionalismo La concepción de que el conocimiento (o algún conocimiento) puede adquirirse de un modo alternativo al uso de los sentidos, mediante el ejercicio de nuestros exclusivos poderes racionales.

Realismo La concepción de que los valores éticos y estéticos, las propiedades matemáticas, etc. existen «ahí fuera» en el mundo, con independencia de nuestro conocimiento o de nuestra experiencia de los mismos.

Reduccionismo Aproximación a un aspecto o ámbito de un discurso que intenta explicar o analizarlo, completa y absolutamente, en otros términos (normalmente más simples o más accesibles), por ejemplo, el fenómeno mental en términos puramente físicos.

Relativismo En ética, la concepción para la cual la corrección o incorrección de las acciones está determinada por, o es relativa a, la cultura y las tradiciones de los grupos o comunidades sociales particulares.

Sintético *véase* Analítico.

Subjetivismo (o antirrealismo) En ética y en estética, la concepción para la cual el fundamento del valor no se encuentra en la realidad externa, sino en nuestras creencias acerca de ella o en nuestras respuestas emocionales a ella.

Utilitarismo En ética, sistema consecuencialista en el que las acciones se juzgan correctas o incorrectas en la medida en que aumenten o disminuyan el bienestar o la «utilidad» humanas; la interpretación clásica de la utilidad es el placer o la felicidad humanos.